PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A01N 43/04, C07H 21/02, 21/04, C12N 5/10, 15/00, 15/09, 15/11, 15/12, 15/63, 15/70, 15/74, 15/79	A1	(11) International Publication Number: WO 98/37764 (43) International Publication Date: 3 September 1998 (03.09.98)

(21) International Application Number: PCT/US98/03895

(22) International Filing Date: 27 February 1998 (27.02.98)

(30) Priority Data:
60/039,388 27 February 1997 (27.02.97) US

(71) Applicants: BAYLOR COLLEGE OF MEDICINE [US/US]; One Baylor Plaza, Houston, TX 77030 (US). JOHNS HOPKINS UNIVERSITY [US/US]; Homewood–Charles & 34th Street, Baltimore, MD 21218 (US). THE UNITED STATES OF AMERICA, represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICESNational Institutes of Health [US/US]; Office of Technology Transfer, Suite 325, 6011 Executive Boulevard, Rockville, MD 20852 (US). THE UNIVERSITY OF UTAH [US/US]; Suite 170, 421 Wakara Way, Salt Lake City, UT 84108 (US).

(72) Inventors: ALLIKMETS, Rando; 5564 Rivendale Place, Frederick, MD 21703 (US). ANDERSON, Kent, L.; P.O. Box 540071, Houston, TX 77254 (US). DEAN, Michael; 1362 Hitchingpost Lane, Frederick, MD 21703 (US). LEPPART, Mark; 1466 Westminster Avenue, Salt Lake City, UT 84105 (US). LEWIS, Richard, A.; 11223 Hermosa Court, Houston, TX 77024 (US). LI, Yixin; Apartment 636, 4045 Linkwood Drive, Houston, TX 77025 (US). LUPSKI, James, R.; 11102 Ashcroft, Houston, TX 77096 (US). NATHANS, Jeremy; 5510 South Bend Road, Baltimore, MD 21209 (US). RATTNER, Amir; Apartment 202, 6503 Copper Ridge Drive, Baltimore, MD 21209 (US). SHROYER, Noah, F.; 4516 1/2 Mt. Vernon Street, Houston, TX 77006 (US). SINGH, Nanda; 1027 Riches Avenue, Salt Lake City, UT 84106 (US). SMALLWOOD, Philip, M.; 5022 Woodbine Road, Woodbine, MD 21797 (US). SUN, Hui; 906A Hooper Avenue, Baltimore, MD 21229 (US).

(74) Agents: CALDWELL, John, W. et al.; Woodcock Washburn Kurtz Mackiewicz & Norris LLP, 46th floor, One Liberty Place, Philadelphia, PA 19103 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: NUCLEIC ACID SEQUENCES FOR ATP-BINDING CASSETTE TRANSPORTER

(57) Abstract

The present invention provides nucleic acid and amino acid sequences of an ATP binding cassette transporter and mutated sequences thereof associated with macular degeneration. Methods of detecting agents that modify ATP–binding cassette transporter comprising combining purified ATP binding cassette transporter and at least one agent suspected of modifying the ATP binding cassette transporter and observing a change in at least one characteristic associated with ATP binding cassette transporter. Methods of detecting macular degeneration is also embodied by the present invention.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

NUCLEIC ACID SEQUENCES FOR ATP-BINDING CASSETTE TRANSPORTER

BACKGROUND OF THE INVENTION

Macular degeneration affects approximately 1.7 million individuals in the U.S. and is the most common cause of acquired visual impairment in those over the age of 65. Stargardt disease (STGD; McKusick Mendelian Inheritance (MIM) #248200) is arguably the most common hereditary recessive macular dystrophy and is characterized by juvenile to young adult onset, central visual impairment, progressive bilateral atrophy of the macular retinal pigment epithelium (RPE) and neuroepithelium, and the frequent appearance of orange-yellow flecks distributed around the macula and/or the midretinal periphery (Stargardt, 1909; Anderson *et al.*, 1995). A clinically similar retinal disorder (Fundus Flavimaculatus, FFM, Franceschetti, 1963) often displays later age of onset and slower progression (Fishman, 1976; Noble and Carr, 1979). From linkage analysis, it has been concluded that STGD and FFM are most likely allelic autosomal recessive disorders with slightly different clinical manifestations caused by mutation(s) of a gene at chromosome 1p13-p21 (Gerber *et al.*, 1995; Anderson *et al.*, 1995). The STGD gene has been localized to a 4 cM region flanked by the

recombinant markers *D1S435* and *D1S236* and a complete yeast artificial chromosome (YAC) contig of the region has been constructed (Anderson *et al.*, 1995). Recently, the location of the STGD/FFM locus on human chromosome 1p has been refined to a 2 cM interval between polymorphic markers *D1S406* and *D1S236* by genetic linkage analysis in an independent set of STGD families (Hoyng *et al.*, 1996). Autosomal dominant disorders with somewhat similar clinical phenotypes to STGD, identified in single large North American pedigrees, have been mapped to chromosome 13q34 (STGD2; MIM #153900; Zhang *et al.*, 1994) and to chromosome 6q11-q14 (STGD3; MIM #600110; Stone *et al.*, 1994), although these conditions are not characterized by the pathognomonic dark choroid observed by fluorescein angiography (Gass, 1987).

Members of the superfamily of mammalian ATP binding cassette (ABC) transporters are being considered as possible candidates for human disease phenotypes. The ABC superfamily includes genes whose products are transmembrane proteins involved in energy-dependent transport of a wide spectrum of substrates across membranes (Childs and Ling, 1994; Dean and Allikmets, 1995). Many disease-causing members of this superfamily result in defects in the transport of specific substrates (CFTR, Riordan *et al.*, 1989; ALD, Mosser *et al.*, 1993; SUR, Thomas *et al.*, 1995; PMP70, Shimozawa *et al.*, 1992; TAP2, de la Salle *et al.*, 1994). In eukaryotes, ABC genes encode typically four domains that include two conserved ATP-binding domains (ATP) and two domains with multiple transmembrane (TM) segments (Hyde *et al.* 1990). The ATP-binding domains of ABC genes contain motifs of characteristic conserved residues (Walker A and B motifs) spaced by 90-120 amino acids. Both this conserved spacing and the "Signature" or "C" motif just upstream of the Walker B site distinguish members of the ABC superfamily from other ATP-binding proteins (Hyde *et al.*, 1990; Michaelis and Berkower, 1995). These features have allowed the isolation of new ABC genes by hybridization, degenerate PCR, and inspection of DNA sequence databases (Allikmets *et al.*, 1993, 1995; Dean *et al.*, 1994; Luciani *et al.*, 1994).

The characterization of twenty-one new members of the ABC superfamily may permit characterization and functions assigned to these genes by determining their map locations and their patterns of expression (Allikmets *et al.*, 1996). That many known ABC genes are involved in inherited human diseases suggests that some of these new loci will also encode proteins mutated in specific genetic disorders. Despite regionally localizing a gene

by mapping, the determination of the precise localization and sequence of one gene nonetheless requires choosing the certain gene from about 250 genes, four to about five million base pairs, from within the regionally localized chromosomal site.

While advancements have been made as described above, mutations in retina-specific ABC transporter (*ABCR*) in patients with recessive macular dystrophy STGD/FFM have not yet been identified to Applicant's knowledge. That *ABCR* expression is limited to photoreceptors, as determined by the present invention, provides evidence as to why *ABCR* has not yet been sequenced. Further, the ABC1 subfamily of ABC transporters is not represented by any homolog in yeast (Michaelis and Berkower, 1995), suggesting that these genes evolved to perform specialized functions in multicellular organisms, which also lends support to why the *ABCR* gene has been difficult to identify. Unlike ABC genes in bacteria, the homologous genes in higher eukaryotes are much less well studied. The fact that prokaryotes contain a large number of ABC genes suggests that many mammalian members of the superfamily remain uncharacterized. The task of studying eukaryote ABC genes is more difficult because of the significantly higher complexity of eukaryotic systems and the apparent difference in function of even highly homologous genes. While ABC proteins are the principal transporters of a number of diverse compounds in bacterial cells, in contrast, eukaryotes have evolved other mechanisms for the transport of many amino acids and sugars. Eukaryotes have other reasons to diversify the role of ABC genes, for example, performing such functions as ion transport, toxin elimination, and secretion of signaling molecules.

Accordingly, there remains a need for the identification of the sequence of the gene, which in mutated forms is associated with retinal and/or macular degenerative diseases, including Stargardt Disease and Fundus Flavimaculatus, for example, in order to provide enhanced diagnoses and improved prognoses and interventional therapies for individuals affected with such diseases.

SUMMARY OF THE INVENTION

The present invention provides sequences encoding an ATP binding cassette transporter. Nucleic acid sequences, including SEQ ID NO: 1 which is a genomic sequence, and SEQ ID NOS: 2 and 5 which are cDNA sequences, are sequences to which the present invention is directed.

A further aspect of the present invention provides ATP binding cassette transporter polypeptides and/or proteins. SEQ ID NOS: 3 and 6 are novel polypeptides of the invention produced from nucleotide sequences encoding the ATP binding cassette transporter. Also within the scope of the present invention is a purified ATP binding cassette transporter.

The present invention also provides an expression vector comprising a nucleic acid sequence encoding an ATP binding cassette transporter, a transformed host cell capable of expressing a nucleic acid sequence encoding an ATP binding cassette transporter, a cell culture capable of expressing an ATP binding cassette transporter, and a protein preparation comprising an ATP binding cassette transporter.

The present invention is also directed to a method of screening for an agent that modifies ATP binding cassette transporter comprising combining purified ATP binding cassette transporter with an agent suspected of modifying ATP binding cassette transporter and observing a change in at least one characteristic associated with ATP binding cassette transporter. The present invention provides methods of identifying an agent that inhibits macular degeneration comprising combining purified ATP binding cassette transporter from a patient suspected of having macular degeneration and an agent suspected interacting with the ATP binding cassette transporter and observing an inhibition in at least one of the characteristics of diseases associated with the ATP binding cassette transporter. In addition, the present invention provides for methods of identifying an agent that induces onset of at least one characteristic associated with ATP binding cassette transporter comprising combining purified wild-type ATP binding cassette transporter with an agent suspected of inducing a macular degenerative disease and observing the onset of a characteristic associated with macular degeneration.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1A and 1B displays the ABCR gene and amplification products. **Figure 1A** displays a physical map of the *ABCR* gene. Mega-YAC clones from the CEPH mega-YAC genomic library (Bellane-Chantelot *et al.*, 1992) encompassing the 4cM critical region for STGD are represented by horizontal bars with shaded circles indicating confirmed positives for STSs by landmark mapping. The individual STS markers and their physical order are shown below the YACs with arrows indicating the centromeric (cen) and telomeric

(1pter) direction (Anderson *et al.*, 1995). The horizontal double head arrow labeled STGD indicates the refined genetic interval delineated by historical recombinants (Anderson *et al.*, 1995). **Figure 1B** displays the results of agarose gel electrophoresis of PCR amplification products with primers from the 5' (GGTCTTCGTGTGTGGTCATT, SEQ ID NO: 114, GGTCCAGTTCTTCCAGAG, SEQ ID NO: 115, labeled 5' ABCR) or 3' (ATCCTCTGACTCAGCAATCACA, SEQ ID NO: 116, TTGCAATTACAAATGCAATGG, SEQ ID NO: 117, labeled 3' ABCR) regions of ABCR on the 13 different YAC DNA templates indicated as diagonals above the gel. The asterisk denotes that YAC 680_b_5 was positive for the 5' ABCR PCR but negative for the 3' ABCR PCR. These data suggest the ABCR gene maps within the interval delineated by markers D1S3361 - D1S236 and is transcribed toward the telomere, as depicted by the open horizontal box.

Figure 2 exhibits the size and tissue distribution of *ABCR* transcripts in the adult rat. A blot of total RNA from the indicated tissues was hybridized with a 1.6 kb mouse *Abcr* probe (top) and a ribosomal protein S26 probe (bottom; Kuwano *et al.*, 1985). The *ABCR* probe revealed a predominant transcript of approximately 8 kb that is found in retina only. The mobility of the 28S and 18S ribosomal RNAs are indicated at the right. B, brain; H, heart; K, kidney; Li, liver; Lu, lung; R, retina; S, spleen.

Figure 3 shows the sequence of the *ABCR* coding region within the genomic *ABCR* sequence, SEQ ID NO: 1. The sequence of the *ABCR* cDNA, SEQ ID NO: 2, is shown with the predicted protein sequence, SEQ ID NO: 3, in one-letter amino acid code below. The location of splice sites is shown by the symbol |.

Figure 4 displays the alignment of the ABCR protein, SEQ ID NO: 3, with other members of the ABC1 subfamily. The deduced amino acid sequence of ABCR is shown aligned to known human and mouse proteins that are members of the same subfamily. Abc1, mouse Abc1, Abc2, mouse Abc2, and ABCC, human ABC gene. The Walker A and B motifs and the Signature motif C are designated by underlining and the letters A, B, and C, respectively.

Figure 5 exhibits the location of *Abcr* from a Jackson BSS Backcross showing a portion of mouse chromosome 3. The map is depicted with the centromere toward the top. A 3 cM scale bar is also shown. Loci mapping to the same position are listed in alphabetical order.

Figure 6 shows the segregation of SSCP variants in exon 49 of the ABCR gene in kindred AR293. Sequence analysis of SSCP bands revealed the existence of wild-type sequence (bands 1 and 3) and mutant sequence (bands 2 and 4). DNA sequencing revealed a 15 base pair deletion, while the affected children (lanes 2 and 3) are homozygous. Haplotype analysis demonstrated homozygosity at the STGD locus in the two affected individuals.

Figure 7A-H shows the localization of *ABCR* transcripts to photoreceptor cells. *In situ* hybridization was performed with digoxygenin-labeled riboprobes and visualized using an alkaline phosphatase conjugated anti-digoxygenin antibody. Figure 7A-D displays hybridization results of retina and choroid from a pigmented mouse (C57/Bl6); Figure 7E and 7F shows hybridization results of retina and choroid from an albino rat; and Figure 7G and 7H exhibits hybridization results of retina from a macaque monkey. Figure 7A, 7E, and 7G display results from a mouse *abcr* antisense probe; Figure 7B exhibit results from a mouse *abcr* sense probe; Figure 7C shows results from a macaque rhodopsin antisense probe; and Figure 7D, 7F, and 7H display results from a mouse blue cone pigment antisense probe. *ABCR* transcripts are localized to the inner segments of the photoreceptor cell layer, a pattern that matches the distribution of rhodopsin transcripts but is distinct from the distribution of cone visual pigment transcripts. Hybridization is not observed in the RPE or choroid, as seen most clearly in the albino rat eye (arrowhead in Figure 7E). The retinal layers indicated in Figure 7B are: OS, outer segments; IS, inner segments; ONL, outer nuclear layer; OPL, outer plexiform layer; INL, inner nuclear layer; IPL, inner plexiform layer; GCL, ganglion cell layer.

Figure 8 provides a pGEM®-T Vector map.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to the nucleic acid and protein sequences encoding ATP binding cassette transporter. The ATP binding cassette transporter of the present invention is retina specific ATP binding cassette transporter (ABCR); more particularly, ABCR may be isolated from retinal cells, preferably photoreceptor cells. The present invention provides nucleotide sequences of *ABCR* including genomic sequences, SEQ ID NO: 1, and cDNA sequences SEQ ID NO: 2 and 5. Novel polypeptide sequences, SEQ

ID NOS: 3 and 6, for ABCR, are the translated products of SEQ ID NOS: 2 and 5, respectively, and are also included in the present invention.

SEQ ID NO:1 provides the human genomic DNA sequence of ABCR. SEQ ID NOS: 2 and 5 provide wild-type cDNA sequences of human ABCR, which result in translated products SEQ ID NOS: 3 and 6, respectively. While not intending to be bound by any particular theory or theories of operation, it is believed that SEQ ID NOS: 2 and 5 are isoforms of ABCR cDNA. The difference between SEQ ID NOS: 2 and 5 may be accounted for by an additional sequence in SEQ ID NO: 2 which is added between bases 4352 and 4353 of SEQ ID NO: 5. This difference is thought to arise from alternative splicing of the nascent transcript of ABCR, in which an alternative exon 30, SEQ ID NO: 4, is excluded. This alternative exon encodes an additional 38 amino acids, SEQ ID NO: 11.

Nucleic acids within in the scope of the present invention include cDNA, RNA, genomic DNA, fragments or portions within the sequences, antisense oligonucleotides. Sequences encoding the ABCR also include amino acid, polypeptide, and protein sequences. Variations in the nucleic acid and polypeptide sequences of the present invention are within the scope of the present invention and include N terminal and C terminal extensions, transcription and translation modifications, and modifications in the cDNA sequence to facilitate and improve transcription and translation efficiency. In addition, changes within the wild-type sequences identified herein which changed sequence retains substantially the same wild-type activity, such that the changed sequences are substantially similar to the ABCR sequences identified, are also considered within the scope of the present invention. Mismatches, insertions, and deletions which permit substantial similarity to the ABCR sequences, such as similarity in residues in hydrophobicity, hydrophilicity, basicity, and acidity, will be known to those of skill in the art once armed with the present disclosure. In addition, the isolated, or purified, sequences of the present invention may be natural, recombinant, synthetic, or a combination thereof. Wild-type activity associated with the ABCR sequences of the present invention include, *inter alia*, all or part of a sequence, or a sequence substantially similar thereto, that codes for ATP binding cassette transporter.

The genomic, SEQ ID NO: 1, and cDNA, SEQ ID NOS: 2 and 5, sequences are identified in Figure 3 and encode ABCR, certain mutations of which are responsible for

the class of retinal disorders known as retinal or macular degenerations. Macular degeneration is characterized by macular dystrophy, various alterations of the peripheral retina, central visual impairment, progressive bilateral atrophy of the macular retinal pigment epithelium (RPE) and neuroepithelium, frequent appearance of orange-yellow flecks distributed around the macula and/or the midretinal periphery, and subretinal deposition of lipofuscin-like material. Retinal and macular degenerative diseases include and are not limited to Stargardt Disease, Fundus Flavimaculatus, age-related macular degeneration, and may include disorders variously called retinitis pigmentosa, combined rod and cone dystrophies, cone dystrophies and degenerations, pattern dystrophy, bull's eye maculopathies, and various other retinal degenerative disorders, some induced by drugs, toxins, environmental influences, and the like. Stargardt Disease is an autosomal recessive retinal disorder characterized by juvenile to adult-onset macular and retinal dystrophy. Fundus Flavimaculatus often displays later age of onset and slower progression. Some environmental insults and drug toxicities may create similar retinal degenerations. Linkage analysis reveals that Stargardt Disease and Fundus Flavimaculatus may be allelic autosomal recessive disorders with slightly different clinical manifestations. The identification of the ABCR gene suggests that different mutations within ABCR may be responsible for these clinical phenomena.

The present invention is also directed to a method of screening for an agent that modifies ATP binding cassette transporter comprising combining purified ATP binding cassette transporter with an agent suspected of modifying ATP binding cassette transporter and observing a change in at least one characteristic associated with ATP binding cassette transporter.

"Modify" and variations thereof include changes such as and not limited to inhibit, suppress, delay, retard, slow, suspend, obstruct, and restrict, as well as induce, encourage, provoke, and cause. Modify may also be defined as complete inhibition such that macular degeneration is arrested, stopped, or blocked. Modifications may, directly or indirectly, inhibit or substantially inhibit, macular degeneration or induce, or substantially induce, macular degeneration, under certain circumstances.

Methods of identifying an agent that inhibits macular degeneration are embodied by the present invention and comprise combining purified ATP binding cassette transporter from a patient suspected of having macular degeneration and an agent suspected

of interacting with the ATP binding cassette transporter and observing an inhibition in at least one of the characteristics of diseases associated with the ATP binding cassette transporter. Accordingly, such methods serve to reduce or prevent macular degeneration, such as in human patients. In addition, the present invention provides for methods of identifying an agent that induces onset of at least one characteristic associated with ATP binding cassette transporter comprising combining purified wild-type ATP binding cassette transporter with an agent suspected of inducing a macular degenerative disease and observing the onset of a characteristic associated with macular degeneration. Thus, such methods provide methods of using laboratory animals to determine causative agents of macular degeneration. The ATP binding cassette transporter may be provided for in the methods identified herein in the form of nucleic acids, such as and not limited to SEQ ID NOS: 1, 2, and 5 or as an amino acid, SEQ ID NOS: 3 and 6, for example. Accordingly, transcription and translation inhibitors may be separately identified. Characteristics associated with macular degeneration include and are not limited to central visual impairment, progressive bilateral atrophy of the macular retinal pigment epithelium (RPE) and neuroepithelium, and the frequent appearance of orange-yellow flecks distributed around the macula and/or the midretinal periphery. Accordingly, observing one or more of the characteristics set forth above results in identification of an agent that induces macular degeneration, whereas reduction or inhibition of at least one of the characteristics results in identification of an agent that inhibits macular degeneration.

Mutational analysis of *ABCR* in Stargardt Disease families revealed thus far seventy four mutations including fifty four single amino acid substitutions, five nonsense mutations resulting in early truncation of the protein, six frame shift mutations resulting in early truncation of the protein, three in-frame deletions resulting in loss of amino acid residues from the protein, and six splice site mutations resulting in incorrect processing of the nascent RNA transcript, see Table 2. Compound heterozygotes for mutations in *ABCR* were found in forty two families. Homozygous mutations were identified in three families with consanguineous parentage. Accordingly, mutations in wild-type *ABCR* which result in activities that are not associated with wild-type *ABCR* are herein referred to as sequences which are associated with macular degeneration. Such mutations include missense mutations, deletions, insertions, substantial differences in hydrophobicity, hydrophilicity, acidity, and

basicity. Characteristics which are associated with retinal or macular degeneration include and are not limited to those characteristics set forth above.

Mutations in wild-type ABCR provide a method of detecting macular degeneration. Retinal or macular degeneration may be detected by obtaining a sample comprising patient nucleic acids from a patient tissue sample; amplifying retina-specific ATP binding cassette receptor specific nucleic acids from the patient nucleic acids to produce a test fragment; obtaining a sample comprising control nucleic acids from a control tissue sample; amplifying control nucleic acids encoding wild-type retina-specific ATP binding cassette receptor to produce a control fragment; comparing the test fragment with the control fragment to detect the presence of a sequence difference in the test fragment, wherein a difference in the test fragment indicates macular degeneration. Mutations in the test fragment, including and not limited to each of the mutations identified above, may provide evidence of macular degeneration.

A purified ABCR protein is also provided by the present invention. The purified ABCR protein may have an amino acid sequence as provided by SEQ ID NOS: 3 and 6.

The present invention is directed to ABCR sequences obtained from mammals from the Order Rodentia, including and not limited to hamsters, rats, and mice; Order Logomorpha, such as rabbits; more particularly the Order Carnivora, including Felines (cats) and Canines (dogs); even more particularly the Order Artiodactyla, Bovines (cows) and Suines (pigs); and the Order Perissodactyla, including Equines (horses); and most particularly the Order Primates, Ceboids and Simoids (monkeys) and Anthropoids (humans and apes). The mammals of most preferred embodiments are humans.

Generally, the sequences of the invention may be produced in host cells transformed with an expression vector comprising a nucleic acid sequence encoding *ABCR*. The transformed cells are cultured under conditions whereby the nucleic acid sequence coding for *ABCR* is expressed. After a suitable amount of time for the protein to accumulate, the protein may be purified from the transformed cells.

A gene coding for *ABCR* may be obtained from a cDNA library. Suitable libraries can be obtained from commercial sources such as Clontech, Palo Alto, CA. Libraries may also be prepared using the following non-limiting examples: hamster insulin-secreting

tumor (HIT), mouse αTC-6, and rat insulinoma (RIN) cells. Positive clones are then subjected to DNA sequencing to determine the presence of a DNA sequence coding for *ABCR*. DNA sequencing is accomplished using the chain termination method of Sanger *et al.*, *Proc. Nat'l. Acad. Sci, U.S.A.*, 1977, *74*, 5463. The DNA sequence encoding *ABCR* is then inserted into an expression vector for later expression in a host cell.

Expression vectors and host cells are selected to form an expression system capable of synthesizing ABCR. Vectors including and not limited to baculovirus vectors may be used in the present invention. Host cells suitable for use in the invention include prokaryotic and eukaryotic cells that can be transformed to stably contain and express ABCR. For example, nucleic acids coding for the recombinant protein may be expressed in prokaryotic or eukaryotic host cells, including the most commonly used bacterial host cell for the production of recombinant proteins, *E. coli*. Other microbial strains may also be used, however, such as *Bacillus subtilis*, and other enterobacteriaceae such as *Salmonella typhimurium* or *Serratia marcescens*, various species of *Pseudomonas*, or other bacterial strains.

The preferable eukaryotic system is yeast, such as *Saccharomyces cerevisiae*. Yeast artificial chromosome (YAC) systems are able to accommodate the large size of ABCR gene sequence or genomic clone. The principle of the YAC system is similar to that used in conventional cloning of DNA. Large fragments of cDNA are ligated into two "arms" of a YAC vector, and the ligation mixture is then introduced into the yeast by transformation. Each of the arms of the YAC vector carries a selectable marker as well as appropriately oriented sequences that function as telomeres in yeast. In addition, one of the two arms carries two small fragments that function as a centromere and as an origin of replication (also called an ARS element-autonomously replicating sequences). Yeast transformants that have taken up and stably maintained an artificial chromosome are identified as colonies on agar plates containing the components necessary for selection of one or both YAC arms. YAC vectors are designed to allow rapid identification of transformants that carry inserts of genomic DNA. Insertion of genomic DNA into the cloning site interrupts a suppressor tRNA gene and results in the formation of red rather than white colonies by yeast strains that carry an amber *ade2* gene.

To clone in YAC vectors, genomic DNA from the test organism is prepared under conditions that result in relatively little shearing such that its average size is several million base pairs. The cDNA is then ligated to the arms of the YAC vector, which has been appropriately prepared to prevent self-ligation. As an alternative to partial digestion with *EcoRI*, YAC vectors may be used that will accept genomic DNA that has been digested to completion with rarely cutting restriction enzymes such as *Not*I or *Mlu*I.

In addition, insect cells, such as *Spodoptera frugiperda*; chicken cells, such as E3C/O and SL-29; mammalian cells, such as HeLa, Chinese hamster ovary cells (CHO), COS-7 or MDCK cells and the like may also be used. The foregoing list is illustrative only and is not intended in any way to limit the types of host cells suitable for expression of the nucleic acid sequences of the invention.

As used herein, expression vectors refer to any type of vector that can be manipulated to contain a nucleic acid sequence coding for *ABCR*, such as plasmid expression vectors, viral vectors, and yeast expression vectors. The selection of the expression vector is based on compatibility with the desired host cell such that expression of the nucleic acid encoding *ABCR* results. Plasmid expression vectors comprise a nucleic acid sequence of the invention operably linked with at least one expression control element such as a promoter. In general, plasmid vectors contain replicon and control sequences derived from species compatible with the host cell. To facilitate selection of plasmids containing nucleic acid sequences of the invention, plasmid vectors may also contain a selectable marker such as a gene coding for antibiotic resistance. Suitable examples include the genes coding for ampicillin, tetracycline, chloramphenicol, or kanamycin resistance.

Suitable expression vectors, promoters, enhancers, and other expression control elements are known in the art and may be found in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, second edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989), incorporated herein by reference in its entirety.

Transformed host cells containing a DNA sequence encoding *ABCR* may then be grown in an appropriate medium for the host. The cells are then grown until product accumulation reaches desired levels at which time the cells are then harvested and the protein product purified in accordance with conventional techniques. Suitable purification methods include, but are not limited to, SDS PAGE electrophoresis, phenylboronate-agarose, reactive

green 19-agarose, concanavalin A sepharose, ion exchange chromatography, affinity chromatography, electrophoresis, dialysis and other methods of purification known in the art.

Protein preparations, of purified or unpurified ABCR by host cells, are accordingly produced which comprise ABCR and other material such as host cell components and/or cell medium, depending on the degree of purification of the protein.

The invention also includes a transgenic non-human animal, including and not limited to mammals, such as and not limited to a mouse, rat, or hamster, comprising a sequence encoding ABCR, or fragment thereof that substantially retains ABCR activity, introduced into the animal or an ancestor of the animal. The sequence may be wild-type or mutant and may be introduced into the animal at the embryonic or adult stage. The sequence is incorporated into the genome of an animal such that it is chromosomally incorporated into an activated state. A transgenic non-human animal has germ cells and somatic cells that contain an ABCR sequence. Embryo cells may be transfected with the gene as it occurs naturally, and transgenic animals are selected in which the gene has integrated into the chromosome at a locus which results in activation. Other activation methods include modifying the gene or its control sequences prior to introduction into the embryo. The embryo may be transfected using a vector containing the gene.

In addition, a transgenic non-human animal may be engineered wherein *ABCR* is suppressed. For purposes of the present invention, suppression of *ABCR* includes, and is not limited to strategies which cause *ABCR* not to be expressed. Such strategies may include and are not limited to inhibition of protein synthesis, pre-mRNA processing, or DNA replication. Each of the above strategies may be accomplished by antisense inhibition of *ABCR* gene expression. Many techniques for transferring antisense sequences into cells are known to those of skill, including and not limited to microinjection, viral-mediated transfer, somatic cell transformation, transgene integration, and the like, as set forth in Pinkert, Carl, *Transgenic Animal Technology*, **1994**, Academic Press, Inc., San Diego, CA, incorporated herein by reference in its entirety.

Further, a transgenic non-human animal may be prepared such that *ABCR* is knocked out. For purposes of the present invention, a knock-out includes and is not limited to disruption or rendering null the *ABCR* gene. A knock-out may be accomplished, for example, with antisense sequences for *ABCR*. The *ABCR* gene may be knocked out by

injection of an antisense sequence for all or part of the *ABCR* sequence such as an antisense sequence for all or part of SEQ ID NO: 2. Once *ABCR* has been rendered null, correlation of the *ABCR* to macular degeneration may be tested. Sequences encoding mutations affecting the *ABCR* may be inserted to test for alterations in various retinal and macular degenerations exhibited by changes in the characteristics associated with retinal and macular degeneration.

An ABCR knock-out may be engineered by inserting synthetic DNA into the animal chromosome by homologous recombination. In this method, sequences flanking the target and insert DNA are identical, allowing strand exchange and crossing over to occur between the target and insert DNA. Sequences to be inserted typically include a gene for a selectable marker, such as drug resistance. Sequences to be targeted are typically coding regions of the genome, in this case part of the ABCR gene. In this process of homologous recombination, targeted sequences are replaced with insert sequences thus disrupting the targeted gene and rendering it nonfunctional. This nonfunctional gene is called a null allele of the gene.

To create the knockout mouse, a DNA construct containing the insert DNA and flanking sequences is made. This DNA construct is transfected into pluripotent embryonic stem cells competent for recombination. The identical flanking sequences align with one another, and chromosomal recombination occurs in which the targeted sequence is replaced with the insert sequence, as described in Bradley, A., Production and Analysis of Chimeric Mice, in *Teratocarcinomas and Embryonic Stem Cells - A Practical Approach*, **1987**, E. Roberson, Editor, IRC Press, pages 113-151. The stem cells are injected into an embryo, which is then implanted into a female animal and allowed to be born. The animals may contain germ cells derived from the injected stem cells, and subsequent matings may produce animals heterozygous and homozygous for the disrupted gene.

Transgenic non-human animals may also be useful for testing nucleic acid changes to identify additional mutations responsible for macular degeneration. A transgenic non-human animal may comprise a recombinant *ABCR*.

The present invention is also directed to gene therapy. For purposes of the present invention, gene therapy refers to the transfer and stable insertion of new genetic information into cells for the therapeutic treatment of diseases or disorders. A foreign sequence or gene is transferred into a cell that proliferates to spread the new sequence or gene

throughout the cell population. Sequences include antisense sequence of all or part of ABCR, such as an antisense sequence to all or part of the sequences identified as SEQ ID NO: 1, 2, and 5. Known methods of gene transfer include microinjection, electroporation, liposomes, chromosome transfer, transfection techniques, calcium-precipitation transfection techniques, and the like. In the instant case, macular degeneration may result from a loss of gene function, as a result of a mutation for example, or a gain of gene function, as a result of an extra copy of a gene, such as three copies of a wild-type gene, or a gene over expressed as a result of a mutation in a promoter, for example. Expression may be altered by activating or deactivating regulatory elements, such as a promoter. A mutation may be corrected by replacing the mutated sequence with a wild-type sequence or inserting an antisense sequence to bind to an over expressed sequence or to a regulatory sequence.

Numerous techniques are known in the art for the introduction of foreign genes into cells and may be used to construct the recombinant cells for purposes of gene therapy, in accordance with this embodiment of the invention. The technique used should provide for the stable transfer of the heterologous gene sequence to the stem cell, so that the heterologous gene sequence is heritable and expressible by stem cell progeny, and so that the necessary development and physiological functions of the recipient cells are not disrupted. Techniques which may be used include but are not limited to chromosome transfer (e.g., cell fusion, chromosome-mediated gene transfer, micro cell-mediated gene transfer), physical methods (e.g., transfection, spheroplast fusion, microinjection, electroporation, liposome carrier), viral vector transfer (e.g., recombinant DNA viruses, recombinant RNA viruses) and the like (described in Cline, M. J., 1985, Pharmac. Ther. 29:69-92, incorporated herein by reference in its entirety).

The term "purified", when used to describe the state of nucleic acid sequences of the invention, refers to nucleic acid sequences substantially free of nucleic acid not coding for *ABCR* or other materials normally associated with nucleic acid in non-recombinant cells, i.e., in its "native state."

The term "purified" or "in purified form" when used to describe the state of an ABCR nucleic acid, protein, polypeptide, or amino acid sequence, refers to sequences substantially free, to at least some degree, of cellular material or other material normally associated with it in its native state. Preferably the sequence has a purity (homogeneity) of

at least about 25% to about 100%, More preferably the purity is at least about 50%, when purified in accordance with standard techniques known in the art.

In accordance with methods of the present invention, methods of detecting retinal or macular degenerations in a patient are provided comprising obtaining a patient tissue sample for testing. The tissue sample may be solid or liquid, a body fluid sample such as and not limited to blood, skin, serum, saliva, sputum, mucus, bone marrow, urine, lymph, and a tear; and feces. In addition, a tissue sample from amniotic fluid or chorion may be provided for the detection of retinal or macular degeneration in utero in accordance with the present invention.

A test fragment is defined herein as an amplified sample comprising *ABCR*-specific nucleic acids from a patient suspected of having retinal or macular degeneration. A control fragment is an amplified sample comprising normal or wild-type *ABCR*-specific nucleic acids from an individual not suspected of having retinal or macular degeneration.

The method of amplifying nucleic acids may be the polymerase chain reaction using a pair of primers wherein at least one primer within the pair is selected from the group consisting of SEQ ID NOS: 12-113. When the polymerase chain reaction is the amplification method of choice, a pair of primers may be used such that one primer of the pair is selected from the group consisting of SEQ ID NOS: 12-113.

Nucleic acids, such as DNA (such as and not limited to genomic DNA and cDNA) and/or RNA (such as and not limited to mRNA), are obtained from the patient sample. Preferably RNA is obtained.

Nucleic acid extraction is followed by amplification of the same by any technique known in the art. The amplification step includes the use of at least one primer sequence which is complementary to a portion of *ABCR*-specific expressed nucleic acids or sequences on flanking intronic genomic sequences in order to amplify exon or coding sequences. Primer sequences useful in the amplification methods include and are not limited to SEQ ID NOS: 12-113, which may be used in the amplification methods. Any primer sequence of about 10 nucleotides to about 35 nucleotides, more preferably about 15 nucleotides to about 30 nucleotides, even more preferably about 17 nucleotides to about 25 nucleotides may be useful in the amplification step of the methods of the present invention. In addition, mismatches within the sequences identified above, which achieve the methods of

the invention, such that the mismatched sequences are substantially complementary and thus hybridizable to the sequence sought to be identified, are also considered within the scope of the disclosure. Mismatches which permit substantial similarity to SEQ ID NOS: 12-113, such as and not limited to sequences with similar hydrophobicity, hydrophilicity, basicity, and acidity, will be known to those of skill in the art once armed with the present disclosure. The primers may also be unmodified or modified. Primers may be prepared by any method known in the art such as by standard phosphoramidite chemistry. See Sambrook *et al.*, *supra.*

The method of amplifying nucleic acids may be the polymerase chain reaction using a pair of primers wherein at least one primer within the pair is selected from the group consisting of SEQ ID NOS: 12-113. When the polymerase chain reaction is the amplification method of choice, a pair of primers may be used such that one primer of the pair is selected from the group consisting of SEQ ID NOS: 12-113.

When an amplification method includes the use of two primers, a first primer and a second primer, such as in the polymerase chain reaction, one of the first primer or second primer may be selected from the group consisting of SEQ ID NOS: 12-113. Any primer pairs which copy and amplify nucleic acids between the pairs pointed toward each other and which are specific for *ABCR* may be used in accordance with the methods of the present invention.

A number of template dependent processes are available to amplify the target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCR) which is described in detail in U.S. Patents 4,683,195, 4,683,202 and 4,800,159, and in Innis *et al.*, *PCR Protocols*, Academic Press, Inc., San Diego CA, 1990, each of which is incorporated herein by reference in its entirety. Briefly, in PCR, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates are added to a reaction mixture along with a DNA polymerase (e.g., *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction products and the process is repeated. Alternatively, a reverse transcriptase PCR

amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies arc well known in the art.

Another method for amplification is the ligase chain reaction (referred to as LCR), disclosed in EPA No. 320,308, incorporated herein by reference in its entirety. In LCR, two complementary probe pairs are prepared, and in the presence of the target sequence, each pair will bind to opposite complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCR, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S. Patent 4,883,750, incorporated herein by reference in its entirety, describes an alternative method of amplification similar to LCR for binding probe pairs to a target sequence.

Qbeta Replicase, described in PCT Application No. PCT/US87/00880, incorporated herein by reference in its entirety, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA which has a region complementary to that of a target is added to a sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence which can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[alpha -thio]triphosphates in one strand of a restriction site (Walker, G. T., *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* **1992**, *89*:392-396, incorporated herein by reference in its entirety), may also be useful in the amplification of nucleic acids in the present invention.

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, i.e. nick translation. A similar method, called Repair Chain Reaction (RCR) is another method of amplification which may be useful in the present invention and which involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA.

ABCR-specific nucleic acids can also be detected using a cyclic probe reaction (CPR). In CPR, a probe having a 3' and 5' sequences of non-*ABCR* specific DNA and middle

sequence of *ABCR* specific RNA is hybridized to DNA which is present in a sample. Upon hybridization, the reaction is treated with RNaseH, and the products of the probe identified as distinctive products, generate a signal which is released after digestion. The original template is annealed to another cycling probe and the reaction is repeated. Thus, CPR involves amplifying a signal generated by hybridization of a probe to a *ABCR*-specific expressed nucleic acid.

Still other amplification methods described in GB Application No. 2 202 328, and in PCT Application No. PCT/US89/01025, each of which is incorporated by reference in its entirety, may be used in accordance with the present invention. In the former application, "modified" primers are used in a PCR like, template and enzyme dependent synthesis. The primers may be modified by labeling with a capture moiety (e.g., biotin) and/or a detector moiety (e.g., enzyme). In the latter application, an excess of labeled probes are added to a sample. In the presence of the target sequence, the probe binds and is cleaved catalytically. After cleavage, the target sequence is released intact to be bound by excess probe. Cleavage of the labeled probe signals the presence of the target sequence.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (Kwoh D., *et al., Proc. Natl. Acad. Sci. (U.S.A.)* **1989**, *86*:1173, Gingeras T. R., *et al.*, PCT Application WO 88/10315, each of which is incorporated herein by reference in its entirety), including nucleic acid sequence based amplification (NASBA) and 3SR. In NASBA, the nucleic acids can be prepared for amplification by standard phenol/chloroform extraction, heat denaturation of a clinical sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride extraction of RNA. These amplification techniques involve annealing a primer which has *ABCR*-specific sequences. Following polymerization, DNA/RNA hybrids are digested with RNase H while double stranded DNA molecules are heat denatured again. In either case the single stranded DNA is made fully double stranded by addition of second *ABCR*-specific primer, followed by polymerization. The double stranded DNA molecules are then multiply transcribed by a polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNAs are reverse transcribed into double stranded DNA, and transcribed once again with a polymerase such as T7 or SP6. The resulting products, whether truncated or complete, indicate *ABCR*-specific sequences.

Davey, C., *et al.*, European Patent Application Publication No. 329,822, incorporated herein by reference in its entirety, disclose a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA ("dsDNA") which may be used in accordance with the present invention. The ssRNA is a first template for a first primer oligonucleotide, which is elongated by reverse transcriptase (RNA-dependent DNA polymerase). The RNA is then removed from resulting DNA:RNA duplex by the action of ribonuclease H (RNase H, an RNase specific for RNA in a duplex with either DNA or RNA). The resultant ssDNA is a second template for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerasc) 5' to its homology to its template. This primer is then extended by DNA polymerase (exemplified by the large "Klenow" fragment of *E. coli* DNA polymerase I), resulting as a double-stranded DNA ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having additionally, at one end, a promoter sequence. This promoter sequence can be used by the appropriate RNA polymerase to make many RNA copies of the DNA. These copies can then re-enter the cycle leading to very swift amplification. With proper choice of enzymes, this amplification can be done isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be chosen to be in the form of either DNA or RNA.

Miller, H. I., *et al.*, PCT application WO 89/06700, incorporated herein by reference in its entirety, disclose a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic; i.e. new templates are not produced from the resultant RNA transcripts. Other amplification methods include "race" disclosed by Frohman, M. A., In: *PCR Protocols: A Guide to Methods and Applications* **1990**, Academic Press, N.Y.) and "one-sided PCR" (Ohara, O., *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* **1989**, *86*:5673-5677), all references herein incorporated by reference in their entirety.

Methods based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying

the di-oligonucleotide (Wu, D. Y. *et al.*, *Genomics* 1989, *4*:560, incorporated herein by reference in its entirety), may also be used in the amplification step of the present invention.

Test fragment and control fragment may be amplified by any amplification methods known to those of skill in the art, including and not limited to the amplification methods set forth above. For purposes of the present invention, amplification of sequences encoding patient and wild-type *ABCR* includes amplification of a portion of a sequence such as and not limited to a portion of an *ABCR* sequence of SEQ ID NO: 1, such as sequence of a length of about 10 nucleotides to about 1,000 nucleotides, more preferably about 10 nucleotides to about 100 nucleotides, or having at least 10 nucleotides occurring anywhere within the SEQ ID NO: 1, where sequence differences are known to occur within *ABCR* test fragments. Thus, for example, a portion of the sequence encoding *ABCR* of a patient sample and a control sample may be amplified to detect sequence differences between these two sequences.

Following amplification of the test fragment and control fragment, comparison between the amplification products of the test fragment and control fragment is carried out. Sequence changes such as and not limited to nucleic acid transition, transversion, and restriction digest pattern alterations may be detected by comparison of the test fragment with the control fragment.

Alternatively, the presence or absence of the amplification product may be detected. The nucleic acids are fragmented into varying sizes of discrete fragments. For example, DNA fragments may be separated according to molecular weight by methods such as and not limited to electrophoresis through an agarose gel matrix. The gels are then analyzed by Southern hybridization. Briefly, DNA in the gel is transferred to a hybridization substrate or matrix such as and not limited to a nitrocellulose sheet and a nylon membrane. A labeled probe encoding an *ABCR* mutation is applied to the matrix under selected hybridization conditions so as to hybridize with complementary DNA localized on the matrix. The probe may be of a length capable of forming a stable duplex. The probe may have a size range of about 200 to about 10,000 nucleotides in length, preferably about 500 nucleotides in length, and more preferably about 2,454 nucleotides in length. Mismatches which permit substantial similarity to the probe, such as and not limited to sequences with similar hydrophobicity, hydrophilicity, basicity, and acidity, will be known to those of skill in the art

once armed with the present disclosure. Various labels for visualization or detection are known to those of skill in the art, such as and not limited to fluorescent staining, ethidium bromide staining for example, avidin/biotin, radioactive labeling such as ^{32}P labeling, and the like. Preferably, the product, such as the PCR product, may be run on an agarose gel and visualized using a stain such as ethidium bromide. See Sambrook *et al.*, *supra*. The matrix may then be analyzed by autoradiography to locate particular fragments which hybridize to the probe. Yet another alternative is the sequencing of the test fragment and the control fragment to identify sequence differences. Methods of nucleic acid sequencing are known to those of skill in the art, including and not limited to the methods of Maxam and Gilbert, *Proc. Natl. Acad. Sci., USA* **1977**, *74*, 560-564 and Sanger, *Proc. Natl. Acad. Sci., USA* **1977**, *74*, 5463-5467.

A pharmaceutical composition comprising all or part of a sequence for ABCR may be delivered to a patient suspected of having retinal or macular degeneration. The sequence may be an antisense sequence. The composition of the present invention may be administered alone or may generally be administered in admixture with a pharmaceutical carrier. The pharmaceutically-acceptable carrier may be selected with regard to the intended route of administration and the standard pharmaceutical practice. The dosage will be about that of the sequence alone and will be set with regard to weight, and clinical condition of the patient. The proportional ratio of active ingredient to carrier will naturally depend, *inter alia*, on the chemical nature, solubility, and stability of the sequence, as well as the dosage contemplated.

The sequences of the invention may be employed in the method of the invention singly or in combination with other compounds, including and not limited to other sequences set forth in the present invention. The method of the invention may also be used in conjunction with other treatments such as and not limited to antibodies, for example. For *in vivo* applications the amount to be administered will also depend on such factors as the age, weight, and clinical condition of the patient. The composition of the present invention may be administered by any suitable route, including as an eye drop, inoculation and injection, for example, intravenous, intraocular, oral, intraperitoneal, intramuscular, subcutaneous, topically, and by absorption through epithelial or mucocutaneous linings, for example, conjunctival, nasal, oral, vaginal, rectal and gastrointestinal.

The mode of administration of the composition may determine the sites in the organism to which the compound will be delivered. For instance, topical application may be administered in creams, ointments, gels, oils, emulsions, pastes, lotions, and the like. For parenteral administration, the composition may be used in the form of sterile aqueous or non-aqueous solution which may contain another solute, for example, sufficient salts, glucose or dextrose to make the solution isotonic. A non-aqueous solution may be comprise an oil, for example. For oral mode of administration, the present invention may be used in the form of tablets, capsules, lozenges, troches, powders, syrups, elixirs, aqueous solutions and suspension, and the like. Various disintegrants, such as starch, and lubricating agents may be used. For oral administration in capsule form, useful diluents are lactose and high molecular weight polyethylene glycols. When aqueous suspensions are required for oral use, certain sweetening and/or flavoring agents may be added.

A diagnostic kit for detecting retinal or macular degeneration comprising in one or more containers at least one primer which is complementary to an *ABCR* sequence and a means for visualizing amplified DNA is also within the scope of the present invention. Alternatively, the kit may comprise two primers. In either case, the primers may be selected from the group consisting of SEQ ID NOS: 12-113, for example. The diagnostic kit may comprise a pair of primers wherein one primer within said pair is complementary to a region of the *ABCR* gene, wherein one of said pair of primers is selected from the group consisting of SEQ ID NO: 12-113, a probe specific to the amplified product, and a means for visualizing amplified DNA, and optionally including one or more size markers, and positive and negative controls. The diagnostic kit of the present invention may comprise one or more of a fluorescent dye such as ethidium bromide stain, ^{32}P, and biotin, as a means for visualizing or detecting amplified DNA. Optionally the kit may include one or more size markers, positive and negative controls, restriction enzymes, and/or a probe specific to the amplified product.

The following examples are illustrative but are not meant to be limiting of the invention.

EXAMPLES:

Identification of the *ABCR* as a Candidate Gene for STGD

One of the 21 new human genes from the ABC superfamily, hereafter called *ABCR* (retina-specific ABC transporter), was identified (Allikmets *et al.* 1996) among expressed sequence tags (ESTs) obtained from 5,000 human retina cDNA clones (Wang, Y., Macke, J.P., Abella, B.S., Andreasson, K., Worley, P., Gilbert, D.J., Copeland, N.G., Jenkins, N.A., and Nathans, J. (1996)) and among ESTs obtained from human retina cDNA clones by the I.M.A.G.E. consortium (Lennon *et al.*, 1996). *ABCR* is closely related to the previously described mouse and human *ABC1* and *ABC2* genes (Luciani *et al.*, 1994; Allikmets *et al.*, 1995). To determine whether *ABCR* might cause a disease, the gene was mapped with a whole genome radiation hybrid panel (GeneBridge 4; Research Genetics, Huntsville, AL). *ABCR* mapped to the human chromosome 1p13-p21 region, close to microsatellite markers *D1S236* and *D1S188*. To define further the location of the gene, PCR primers, 3'UTR-For 5'ATCCTCTGACTCAGCAATCACA, SEQ ID NO: 7, and 3'UTR-Rev 5'TTGCAATTACAAATGCAATGG, SEQ ID NO: 8, from the putative 3' untranslated region were used to screen YACs from the previously described contig between these anonymous markers (Anderson *et al.*, 1995). At least 12 YACs contain the 3' end of the *ABCR* gene, including 924_e_9, 759_d_7, 775_c_2, 782_b_4, 982_g_5, 775_b_2, 765_a_3, 751_f_2, 848_e_3, 943_h_8, 934_g_7, and 944_b_12 (Figure 1). These YACs delineate a region containing the STGD gene between markers *D1S3361* and *D1S236* (Anderson *et al.*, 1995).

Expression of the *ABCR* Gene

Additional support suggesting that *ABCR* is a candidate STGD gene came from expression studies and inspection of the EST databases.

Searches of the dbEST (Boguski *et al.*, 1993) database were performed with BLAST on the NCBI file server (Altschul *et al.*, 1990). Amino acid alignments were generated with PILEUP (Feng and Doolittle, 1987). Sequences were analyzed with programs of the Genetics Computer Group package (Devereaux *et al.*, 1984) on a VAX computer.

Clones corresponding to the mouse ortholog of the human *ABCR* gene were isolated from the mouse retina cDNA library and end-sequenced. The chromosomal location of the mouse *ABCR* gene was determined on The Jackson Laboratory (Bar Harbor, ME)

interspecific backcross mapping panel (C57BL/6JEi X SPRET/Ei)F1 X SPRET/Ei (Rowe *et al.*, 1994) known as Jackson BSS. Mapping was performed by SSCP analysis with the primers MABCR1F 5'ATC CAT ACC CTT CCC ACT CC, SEQ ID NO: 9, and MABCR1R 5' GCA GCA GAA GAT AAG CAC ACC, SEQ ID NO. 10. The allele pattern of the *Abcr* was compared to the 250 other loci mapped previously in the Jackson BSS cross (http://www.jax.org).

DNA fragments used as probes were purified on a 1% low-melting temperature agarose gel. The probe sequences are set forth within the genomic sequence of SEQ ID NO: 1 and Figure 3. DNA was labeled directly in agarose with the Random Primed DNA Labeling Kit (Boehringer Mannheim, Indianapolis, IN) and hybridized to multiple tissue Northern blot and a Master blot (Clontech, Palo Alto, CA), according to the manufacturer's instructions. Each blot contained 2 μg of poly A^+ RNA from various human tissues. Total RNA was isolated from adult rat tissues using the guanidinium thiocyanate method (Chomczynski and Saachi, 1987) and resolved by agarose gel electrophoresis in the presence of formaldehyde (Sambrook *et al.*, 1989). Hybridization with the mouse *ABCR* probe was performed in 50% formamide, 5X SSC at 42°C, and filters were washed in 0.1X SSC at 68°C.

Hybridization of a 3' *ABCR* cDNA probe to a multiple tissue Northern blot and a MasterBlot (Clontech, Palo Alto, CA) indicated that the gene was not expressed detectably in any of the 50 non-retinal fetal and adult tissues examined, consistent with the observation that all 12 of the *ABCR* clones in the EST database originated from retinal cDNA libraries. Furthermore, screening cDNA libraries from both developing mouse eye and adult human retina with *ABCR* probes revealed an estimated at 0.1%-1% frequency of *ABCR* clones of all cDNA clones in the library. Hybridization of the *ABCR* probe to a Northern blot containing total RNA from rat retina and other tissues showed that the expression of this gene is uniquely retina-specific (Figure 2). The transcript size is estimated to be 8 kb.

Sequence and Exon/Intron Structure of the *ABCR* cDNA

Several ESTs that were derived from retina cDNA libraries and had high similarity to the mouse *Abc1* gene were used to facilitate the assembly of most of the *ABCR* cDNA sequence. Retina cDNA clones were linked by RT-PCR, and repetitive screening of

a human retina cDNA library with 3' and 5' PCR probes together with 5' RACE were used to characterize the terminal sequences of the gene.

cDNA clones containing *ABCR* sequences were obtained from a human retina cDNA library (Nathans *et al.*, 1986) and sequenced fully. Primers were designed from the sequences of cDNA clones from 5' and 3' regions of the gene and used to link the identified cDNA clones by RT-PCR with retina QUICK-Clone cDNA (Clontech, Palo Alto, CA) as a template. PCR products were cloned into pGEM®-T vector (Promega, Madison, WI). Mouse *ABCR* cDNA clones were obtained from screening a developing mouse eye cDNA library (H. Sun, A. Lanahan, and J. Nathans, unpublished). The pGEM®-T Vector is prepared by cutting pGEM®-5Zf(+) DNA with EcoR V and adding to a 3' terminal thymidine to both ends. These single 3'-T overhangs at the insertion site greatly improve the efficiency of ligation of PCR products because of the nontemplate-dependent addition of a single deoxyadenosine (A) to the 3'-ends of PCR products by many thermostable polymerases. The pGEM®-5Zf(+) Vector contains the origin of replication of the filamentous phage f1 and can be used to produce ssDNA. The plasmid also contains T7 and SP6 RNA polymerase promoters flanking a multiple cloning region within the α-peptide coding region for the enzyme β-galactosidase. Insertional inactivation of the α-peptide allows recombinant clones to be identified directly by color screening on indicator plates. cDNA clones from various regions of the *ABCR* gene were used as probes to screen a human genomic library in Lambda FIX II (#946203, Stratagene, LaJolla, CA). Overlapping phage clones were mapped by *Eco*RI and *Bam*HI digestion. A total of 6.9 kb of the *ABCR* sequence was assembled, (Figure 3) resulting in a 6540 bp (2180 amino acid) open reading frame.

Screening of a bacteriophage lambda human genomic library with cDNA probes yielded a contig that spans approximately 100 kb and contains the majority of the *ABCR* coding region. The exon/intron structure of all fifty one exons of the gene were characterized by direct sequencing of genomic and cDNA clones. Intron sizes were estimated from the sizes of PCR products using primers from adjacent exons with genomic phage clones as templates.

Primers for the cDNA sequences of the *ABCR* were designed with the PRIMER program (Lincoln *et al.*, 1991). Both *ABCR* cDNA clones and genomic clones became templates for sequencing. Sequencing was performed with the Taq Dyedeoxy Terminator

Cycle Sequencing kit (Applied Biosystems, Foster City, CA), according to the manufacturer's instructions. Sequencing reactions were resolved on an ABI 373A automated sequencer. Positions of introns were determined by comparison between genomic and cDNA sequences. Primers for amplification of individual exons were designed from adjacent intron sequences 20-50 bp from the splice site and are set forth in Table 1.

Table 1 Exon/intron Primers for ABCR

PRIMER	SEQUENCE	SEQ ID NO
ABCR.EXON1:F	ACCCTCTGCTAAGCTCAGAG	12
ABCR.EXON1:R	ACCCCACACTTCCAACCTG	13
ABCR.EXON2:F	AAGTCCTACTGCACACATGG	14
ABCR.EXON2:R	ACACTCCCACCCCAAGATC	15
ABCR.EXON3:F	TTCCCAAAAAGGCCAACTC	16
ABCR.EXON3:R	CACGCACGTGTGCATTCAG	17
ABCR.EXON4:F	GCTATTTCCTTATTAATGAGGC	18
ABCR.EXON4:R	CCAACTCTCCCTGTTCTTTC	19
ABCR.EXON5:F	TGTTTCCAATCGACTCTGGC	20
ABCR.EXON5:R	TTCTTGCCTTTCTCAGGCTGG	21
ABCR.EXON6:F	GTATTCCCAGGTTCTGTGG	22
ABCR.EXON6:R	TACCCCAGGAATCACCTTG	23
ABCR.EXON7:F	AGCATATAGGAGATCAGACTG	24
ABCR.EXON7:R	TGACATAAGTGGGGTAAATGG	25
ABCR.EXON8:F	GAGCATTGGCCTCACAGCAG	26
ABCR.EXON8:R	CCCCAGGTTTGTTTCACC	27

Table 1 Exon/intron Primers for ABCR (continued)

PRIMER	SEQUENCE	SEQ ID NO
ABCR.EXON9:F	AGACATGTGATGTGGATACAC	28
ABCR.EXON9:R	GTGGGAGGTCCAGGGTACAC	29
ABCR.EXON10:F	AGGGGCAGAAAAGACACAC	30
ABCR.EXON10:R	TAGCGATTAACTCTTTCCTGG	31
ABCR.EXON11:F	CTCTTCAGGGAGCCTTAGC	32
ABCR.EXON11:R	TTCAAGACCACTTGACTTGC	33
ABCR.EXON12:F	TGGGACAGCAGCCTTATC	34
ABCR.EXON12:R	CCAAATGTAATTTCCCACTGAC	35
ABCR.EXON13:F	AATGAGTTCCGAGTCACCCTG	36
ABCR.EXON13:R	CCCATTCGCGTGTCATGG	37
ABCR.EXON14:F	TCCATCTGGGCTTTGTTCTC	38
ABCR.EXON14:R	AATCCAGGCACATGAACAGG	39
ABCR.EXON15:F	AGGCTGGTGGGAGAGAGC	40
ABCR.EXON15:R	AGTGGACCCCCTCAGAGG	41
ABCR.EXON16:F	CTGTTGCATTGGATAAAAGGC	42
ABCR.EXON16:R	GATGAATGGAGAGGGCTGG	43
ABCR.EXON17:F	CTGCGGTAAGGTAGGATAGGG	44
ABCR.EXON17:R	CACACCGTTTACATAGAGGGC	45
ABCR.EXON18:F	CCTCTCCCCTCCTTTCCTG	46
ABCR.EXON18:R	GTCAGTTTCCGTAGGCTTC	47

Table 1 Exon/intron Primers for ABCR (continued)

PRIMER	SEQUENCE	SEQ ID NO
ABCR.EXON19:F	TGGGGCCATGTAATTAGGC	48
ABCR.EXON19:R	TGGGAAAGAGTAGACAGCCG	49
ABCR.EXON20:F	ACTGAACCTGGTGTGGGG	50
ABCR.EXON20:R	TATCTCTGCCTGTGCCCAG	51
ABCR.EXON21:F	GTAAGATCAGCTGCTGGAAG	52
ABCR.EXON21:R	GAAGCTCTCCTGCACCAAGC	53
ABCR.EXON22:F	AGGTACCCCCACAATGCC	54
ABCR.EXON22:R	TCATTGTGGTTCCAGTACTCAG	55
ABCR.EXON23:F	TTTTTGCAACTATATAGCCAGG	56
ABCR.EXON23:R	AGCCTGTGTGAGTAGCCATG	57
ABCR.EXON24:F	GCATCAGGGCGAGGCTGTC	58
ABCR.EXON24:R	CCCAGCAATACTGGGAGATG	59
ABCR.EXON25:F	GGTAACCTCACAGTCTTCC	60
ABCR.EXON25:R	GGGAACGATGGCTTTTTGC	61
ABCR.EXON26:F	TCCCATTATGAAGCAATACC	62
ABCR.EXON26:R	CCTTAGACTTTCGAGATGG	63
ABCR.EXON27:F	GCTACCAGCCTGGTATTTCATTG	64
ABCR.EXON27:R	GTTATAACCCATGCCTGAAG	65

Table 1 Exon/intron Primers for ABCR (continued)

PRIMER	SEQUENCE	SEQ ID NO
ABCR.EXON28:F	TGCACGCGCACGTGTGAC	66
ABCR.EXON28:R	TGAAGGTCCCAGTGAAGTGGG	67
ABCR.EXON29:F	CAGCAGCTATCCAGTAAAGG	68
ABCR.EXON29:R	AACGCCTGCCATCTTGAAC	69
ABCR.EXON30:F	GTTGGGCACAATTCTTATGC	70
ABCR.EXON30:R	GTTGTTTGGAGGTCAGGTAC	71
ABCR.EXON31:F	AACATCACCCAGCTGTTCCAG	72
ABCR.EXON31:R	ACTCAGGAGATACCAGGGAC	73
ABCR.EXON32:F	GGAAGACAACAAGCAGTTTCAC	74
ABCR.EXON32:R	ATCTACTGCCCTGATCATAC	75
ABCR.EXON33:F	AAGACTGAGACTTCAGTCTTC	76
ABCR.EXON33:R	GGTGTGCCTTTTAAAAGTGTGC	77
ABCR.EXON34:F	TTCATGTTTCCCTACAAAACCC	78
ABCR.EXON34:R	CATGAGAGTTTCTCATTCATGG	79
ABCR.EXON35:F	TGTTTACATGGTTTTTAGGGCC	80
ABCR.EXON35:R	TTCAGCAGGAGGAGGGATG	81
ABCR.EXON36:F	CCTTTCCTTCACTGATTTCTGC	82
ABCR.EXON36:R	AATCAGCACTTCGCGGTG	83

Table 1 Exon/intron Primers for ABCR (continued)

PRIMER	SEQUENCE	SEQ ID NO
ABCR.EXON37:F	TGTAAGGCCTTCCCAAAGC	84
ABCR.EXON37:R	TGGTCCTTCAGCGCACACAC	85
ABCR.EXON38:F	CATTTTGCAGAGCTGGCAGC	86
ABCR.EXON38:R	CTTCTGTCAGGAGATGATCC	87
ABCR.EXON39:F	GGAGTGCATTATATCCAGACG	88
ABCR.EXON39:R	CCTGGCTCTGCTTGACCAAC	89
ABCR.EXON40:F	TGCTGTCCTGTGAGAGCATC	90
ABCR.EXON40:R	GTAACCCTCCCAGCTTTGG	91
ABCR.EXON41:F	CAGTTCCCACATAAGGCCTG	92
ABCR.EXON41:R	CAGTTCTGGATGCCCTGAG	93
ABCR.EXON42:F	GAAGAGAGGTCCCATGGAAAGG	94
ABCR.EXON42:R	GCTTGCATAAGCATATCAATTG	95
ABCR.EXON43:F	CTCCTAAACCATCCTTTGCTC	96
ABCR.EXON43:R	AGGCAGGCACAAGAGCTG	97
ABCR.EXON44:F	CTTACCCTGGGGCCTGAC	98
ABCR.EXON44:R	CTCAGAGCCACCCTACTATAG	99
ABCR.EXON45:F	GAAGCTTCTCCAGCCCTAGC	100
ABCR.EXON45:R	TGCACTCTCATGAAACAGGC	101

Table 1 Exon/intron Primers for ABCR (continued)

PRIMER	SEQUENCE	SEQ ID NO
ABCR.EXON46:F	GTTTGGGGTGTTTGCTTGTC	102
ABCR.EXON46:R	ACCTCTTTCCCCAACCCAGAG	103
ABCR.EXON47:F	GAAGCAGTAATCAGAAGGGC	104
ABCR.EXON47:R	GCCTCACATTCTTCCATGCTG	105
ABCR.EXON48:F	TCACATCCCACAGGCAAGAG	106
ABCR.EXON48:R	TTCCAAGTGTCAATGGAGAAC	107
ABCR.EXON49:F	ATTACCTTAGGCCCAACCAC	108
ABCR.EXON49:R	ACACTGGGTGTTCTGGACC	109
ABCR.EXON50:F	GTGTAGGGTGGTGTTTTCC	110
ABCR.EXON50:R	AAGCCCAGTGAACCAGCTGG	111
ABCR.EXON51:F	TCAGCTGAGTGCCCTTCAG	112
ABCR.EXON51:R	AGGTGAGCAAGTCAGTTTCGG	113

In Table 1, "F" indicates forward, i.e., 5' to 3', "R" indicates reverse, i.e., 3' to 5'. PCR conditions were 95°C for 8 minutes; 5 cycles at 62°C for 20 seconds, 72°C for 30 seconds; 35 cycles at 60°C for 20 seconds, 72°C for 30 seconds; 72°C for 5 minutes (except that [a] was performed at 94° C for 5 minutes); 5 cycles at 94° C for 40 seconds; 60° C for 30 seconds; 72° C for 20 seconds; 35 cycles at 94° C for 40 seconds; 56° C for 30 seconds; 72° C for 20 seconds, and 72° C for 5 minutes.

Amplification of exons was performed with AmpliTaq Gold polymerase in a 25 μl volume in 1X PCR buffer supplied by the manufacturer (Perkin Elmer, Foster City, CA). Samples were heated to 95°C for 10 minutes and amplified for 35-40 cycles at 96°C for 20 seconds; 58°C for 30 seconds; and 72°C for 30 seconds. PCR products were analyzed

on 1-1.5% agarose gels and in some cases digested with an appropriate restriction enzymes to verify their sequence. Primer sequences and specific reaction conditions are set forth in Table 1. The sequence of the *ABCR* cDNA has been deposited with GenBank under accession # U88667.

Homology to ABC Superfamily Members

A BLAST search revealed that *ABCR* is most closely related to the previously characterized mouse *Abc1* and *Abc2* genes (Luciani *et al.*, 1994) and to another human gene (*ABCC*) which maps to chromosome 16p13.3 (Klugbauer and Hofmann, 1996). These genes, together with *ABCR* and a gene from *C. elegans* (GenBank #Z29117), form a subfamily of genes specific to multicellular organisms and not represented in yeast (Michaelis and Berkower, 1995; Allikmets *et al.*, 1996). Alignment of the cDNA sequence of *ABCR* with the *Abc1*, *Abc2*, and *ABCC* genes revealed, as expected, the highest degree of homology within the ATP-binding cassettes. The predicted amino acid identity of the *ABCR* gene to mouse *Abc1* was 70% within the ATP-binding domains; even within hydrophobic membrane-spanning segments, homology ranged between 55 and 85% (Figure 4). The putative *ABCR* initiator methionine shown in Figures 3 and 4 corresponds to a methionine codon at the 5' end of *Abc1* (Luciani *et al.*, 1994).

ABCR shows the composition of a typical full-length ABC transporter that consists of two transmembrane domains (TM), each with six membrane spanning hydrophobic segments, as predicted by a hydropathy plot (data not shown), and two highly conserved ATP-binding domains (Figures 3 and 4). In addition, the HH1 hydrophobic domain, located between the first ATP and second TM domain and specific to this subfamily (Luciani *et al.*, 1994), showed a predicted 57% amino acid identity (24 of 42 amino acids) with the mouse *Abc1* gene.

To characterize the mouse ortholog of *ABCR*, cDNA clones from a developing mouse eye library were isolated. A partial sequence of the mouse cDNA was utilized to design PCR primers to map the mouse *Abcr* gene in an interspecific backcross mapping panel (Jackson BSS). The allele pattern of *Abcr* was compared to 2450 other loci mapped previously in the Jackson BSS cross; linkage was found to the distal end of chromosome 3 (Figure 5). No recombinants were observed between *Abcr* and *D13Mit13*. This region of the

mouse genome is syntenic with human chromosome 1p13-p21. Thus far, no eye disease phenotype has been mapped to this region of mouse chromosome 3.

Compound Heterozygous and Homozygous Mutations in STGD Patients

One hundred forty-five North American and three Saudi Arabian families with STGD/FFM were examined. Among these, at least four were consanguineous families in which the parents were first cousins. Entry criteria for the characterization of the clinical and angiographic diagnosis of Stargardt disease, ascertainment of the families, and methodology for their collection, including the consanguineous families from Saudi Arabia, were as provided in Anderson *et al.*, 1995; and Anderson, 1996.

Mutational analysis of the *ABCR* gene was pursued in the above identified one hundred forty-eight STGD families previously ascertained by strict definitional criteria and shown to be linked to chromosome 1p (Anderson *et al.*, 1995; Anderson, 1996). To date, all 51 exons have been used for mutation analysis.

Mutations were detected by a combined SSCP (Orita *et al.*, 1989) and heteroduplex analysis (White *et al.*, 1992) under optimized conditions (Glavač and Dean, 1993). Genomic DNA samples (50 ng) were amplified with AmpliTaq Gold polymerase in 1X PCR buffer supplied by the manufacturer (Perkin Elmer, Foster City, CA) containing [α-^{32}P] dCTP. Samples were heated to 95°C for 10 minutes and amplified for 35-40 cycles at 96°C for 20 seconds; 58°C for 30 seconds; and 72°C for 30 seconds. Products were diluted in 1:3 stop solution, denatured at 95°C for 5 minutes, chilled in ice for 5 minutes, and loaded on gels. Gel formulations include 6% acrylamide:Bis (2.6% cross-linking), 10% glycerol at room temperature, 12W; and 10% acrylamide:Bis (1.5% cross-linking), at 4°C, 70W. Gels were run for 2-16 hours (3000 Vh/100 bp), dried, and exposed to X-ray film for 2-12 hours. Some exons were analyzed by SSCP with MDE acrylamide (FMC Bioproducts, Rockland, ME) with and without 10% glycerol for 18 hours, 4 watts at room temperature with α-P^{32}-dCTP labeled DNA. Heteroduplexes were identified from the double-stranded DNA at the bottom of the gels, and SSCPs were identified from the single-stranded region. Samples showing variation were compared with other family members to assess segregation of the alleles and with at least 40 unrelated control samples, from either Caucasian or Saudi Arabian populations, to distinguish mutations from polymorphisms unrelated to STGD. PCR products

with SSCP or heteroduplex variants were obtained in a 25 µl volume, separated on a 1% agarose gel, and isolated by a DNA purification kit (PGC Scientific, Frederick, MD). Sequencing was performed on an ABI sequencer with both dye primer and dye terminator chemistry.

Some mutations were identified with a heteroduplex analysis protocol (Roa *et al.*, 1993). Equimolar amounts of control and patient PCR products were mixed in 0.2 ml tubes. Two volumes of PCR product from a normal individual served as a negative control, and *MPZ* exon 3 from patient BAB731 as a positive control (Roa *et al.*, 1996). Samples were denatured at 95°C for 2 minutes and cooled to 35°C at a rate of 1°C/minute. Samples were loaded onto 1.0 mm thick, 40 cm MDE gels (FMC Bioproducts, Rockland, ME), electrophoresed at 600-800 V for 15-20 hours, and visualized with ethidium bromide. Samples showing a variant band were reamplified with biotinylated forward and reverse primers and immobilized on streptavidin-conjugated beads (Warner *et al.* 1996). The resulting single strands were sequenced by the dideoxy-sequencing method with Sequenase 2.0 (Amersham, Arlington Heights, IL).

A total of seventy five mutations were identified, the majority representing missense mutations in conserved amino acid positions. However, several insertions and deletions representing frameshifts were also found (Table 2). The sequence of two mutations are shown in Figure 6A and 6B. Two missense alterations (D847H, R943Q) were found in at least one control individual, suggesting that they are neutral polymorphisms. The remaining mutations were found in patients having macular degeneration and were not found in at least 220 unrelated normal controls (440 chromosomes), consistent with the interpretation that these alterations represent disease-causing mutations, not polymorphisms. One of the mutations, 5892+1 G→T, occurs in family AR144 in which one of the affected children is recombinant for the flanking marker *D1S236* (Anderson *et al.*, 1995). This mutation, however, is present in the father as well as in both affected children. Therefore, the *ABCR* gene is non-recombinant with respect to the Stargardt disease locus.

The mutations are scattered throughout the coding sequence of the *ABCR* gene (see Table 2 and Figure 3), although clustering within the conserved regions of the ATP-binding domains is noticeable. Homozygous mutations were detected in three likely consanguineous families, two Saudi Arabian and one North American (Anderson *et al.*, 1995),

in each of which only the affected individuals inherited the identical disease allele (Table 2; Figure 6C). Forty two compound heterozygous families were identified in which the two disease alleles were transmitted from different parents to only the affected offspring (Table 2).

Table 2. Mutations in the *ABCR* gene in STGD Families

Nucleotide	Amino Acid	#Families	Exon
0223T->G	C75G	1	3
0634C->T	R212C	1	6
0664del13	fs	1	6
0746A->G	D249G	1	6
1018T->G	Y340D	2	8
1411G->A	E471K	1	11
1569T->G	D523E	1	12
1715G->A	R572Q	2	12
1715G->C	R572P	1	12
1804C->T	R602W	1	13
1822T->A	F608I	1	13
1917C->A	Y639X	1	13
2453G->A	G818E	1	16
2461T->A	W821R	1	16
2536G->C	D846H	1	16
2588G->C	G863A	11	17
2791G->A	V931M	1	19
2827C->T	R943W	1	19
2884delC	fs	1	19
2894A->G	N965S	3	19
3083C->T	A1028V	14	21
3211delGT	fs	1	22
3212C->T	S1071L	1	22
3215T->C	V1072A	1	22
3259G->A	E1087K	1	22
3322C->T	R1108C	6	22
3364G->A	E1122K	1	23
3385G->T	R1129C	1	23
3386G->T	R1129L	1	23
3602T->G	L1201R	1	24
3610G->A	D1204N	1	25
4139C->T	P1380L	2	28
4195G->A	E1399K	1	28
4222T->C	W1408R	3	28
4232insTATG	fs	1	28
4253+5G->T	splice	1	28
4297G->A	V1433I	1	29
4316G->A	G1439D	1	29

Table 2. Mutations in the *ABCR* gene in STGD Families

Nucleotide	Amino Acid	#Families	Exon
4319T->C	F1440S	1	29
4346G->A	W1449X	1	29
4462T->C	C1488R	1	30
4469G->A	C1490Y	1	31
4577C->T	T1526M	6	32
4594G->A	D1532N	2	32
4947delC	fs	1	36
5041del15	VVAIC1681del	1	37
5196+2T->C	splice	1	37
5281del9	PAL1761del	1	38
5459G->C	R1820P	1	39
5512C->T	H1838Y	1	40
5527C->T	R1843W	1	40
5585+1G->A	splice	1	41
5657G->A	G1886E	1	41
5693G->A	R1898H	4	41
5714+5G->A	splice	8	41
5882G->A	G1961E	16	43
5898+1G->A	splice	3	43
5908C->T	L1970F	1	44
5929G->A	G1977S	1	44
6005+1G->T	splice	1	44
6079C->T	L2027F	11	45
6088C->T	R2030X	1	45
6089G->A	R2030Q	1	45
6112C->T	R2038W	1	45
6148G->C	V2050L	2	46
6166A->T	K2056X	1	46
6229C->T	R2077W	1	46
6286G->A	E2096K	1	47
6316C->T	R2106C	1	47
6391G->A	E2131K	1	48
6415C->T	R2139W	1	48
6445C->T	R2149X	1	48
6543del36	1181del12	1	49
6709delG	fs	1	49

Mutations are named according to standard nomenclature. The column headed "Exon" denotes which of the 51 exons of ABCR contain the mutation. The column headed "# Families" denotes the number of Stargardt families which displayed the mutation. The column headed "Nucleotide" gives the base number starting from the A in the initiator ATG,

followed by the wild type sequence and an arrow indicating the base it is changed to; del indicates a deletion of selected bases at the given position in the ABCR gene; ins indicates an insertion of selected bases at the given position; splice donor site mutations are indicated by the number of the last base of the given exon, followed by a plus sign and the number of bases into the intron where the mutation occurs. The column headed "Amino Acid" denotes the amino acid change a given mutation causes; fs indicates a frameshift mutation leading to a truncated protein; splice indicates a splice donor site mutation; del indicates an in-frame deletion of the given amino acids.

Mutations are named according to standard nomenclature. Exon numbering according to the nucleotide position starting from the A in the initiator ATG.

***In Situ* Hybridization**

STGD is characterized histologically by a massive accumulation of a lipofuscin-like substance in the retinal pigment epithelium (RPE). This characteristic has led to the suggestion that STGD represents an RPE storage disorder (Blacharski *et al.*, 1988). It was therefore of interest that *ABCR* transcripts were found to be abundant in the retina. To identify the site(s) of *ABCR* gene expression at higher resolution and to determine whether the gene is also expressed in the RPE, the distribution of *ABCR* transcripts was visualized by *in situ* hybridization to mouse, rat, bovine, and macaque ocular tissues.

In situ hybridization with digoxigenin-labeled riboprobes was performed as described by Schaeren-Wiemers and Gerfin-Moser, 1993. For mouse and rat, unfixed whole eyes were frozen and sectioned; macaque retinas were obtained following cardiac perfusion with paraformaldehyde as described (Zhou *et al.*, 1996). An extra incubation of 30 min in 1% Triton X-100, 1X PBS was applied to the fixed monkey retina sections immediately after the acetylation step. The templates for probe synthesis were: (1) a 1.6 kb fragment encompassing the 3' end of the mouse *Abcr* coding region, (2) a full length cDNA clone encoding the mouse blue cone pigment (Chiu *et al.*, 1994), and (3) a macaque rhodopsin coding region segment encoding residues 133 to 254 (Nickells, R. W., Burgoyne, C.F., Quigley, H.A., and Zack, D.J. (1995)).

This analysis showed that *ABCR* transcripts are present exclusively within photoreceptor cells (Figure 7). *ABCR* transcripts are localized principally to the rod inner

segments, a distribution that closely matches that of rhodopsin gene transcripts. Interestingly, *ABCR* hybridization was not observed at detectable levels in cone photoreceptors, as judged by comparisons with the hybridization patterns obtained with a blue cone pigment probe (compare Figure 7A and Figure 7D, Figure 7E with Figure 7F and Figure 7G with Figure 7H). Because melanin granules might obscure a weak hybridization signal in the RPE of a pigmented animal, the distribution of *ABCR* transcripts was also examined in both albino rats and albino mice. In these experiments, the *ABCR* hybridization signal was seen in the photoreceptor inner segments and was unequivocally absent from the RPE (Figure 7E). Given that *ABCR* transcripts in each of these mammals, including a primate, are photoreceptor-specific, it is highly likely that the distribution of *ABCR* transcripts conforms to this pattern as well in the human retina.

The disclosures of each patent, patent application and publication cited or described in this document are hereby incorporated herein by reference, in their entirety.

Various modifications of the invention in addition to those shown and described herein will be apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims.

References

Allikmets, R., Singh, N., Sun, H., Shroyer, N.F., Hutchinson, A., Chidambaram, A., Gerrard, B., Baird, L., Stauffer, D., Peiffer, A., Rattner, A., Smallwood, P., Li, Y., Anderson, K.L., Lewis, R.A., Nathans, J., Leppert, M., Dean, M., Lupski, J.R., (1997) A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is mutated in recessive Stargardt macular dystrophy. Nature Genetics. 15(3):236-46.

Allikmets, R., Shroyer, N.F., Singh, N., Seddon, J.M., Lewis, R.A.,Bernsteinm P.S., Peiffer, A., Zabriskie, N.A., Li, Y., Hutchinson, A., Dean, M., Lupski, J.R., Leppert, M., (1997) Mutation of the Stargardt disease gene (ABCR) in age-related macular degeneration. Science. 277(5333):1805-7.

Allikmets, R., Gerrard, B., Hutchinson, A., and Dean, M. (1996). Characterization of the human ABC superfamily: Isolation and mapping of 21 new genes using the Expressed Sequence Tag database. Hum. Mol. Genet. *5*, 1649-1655.

Allikmets, R., Gerrard, B., Glavač, D., Ravnik-Glavač, M., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., Modi, W., and Dean, M. (1995). Characterization and mapping of three new mammalian ATP-binding transporter genes from an EST database. Mam. Genome *6*, 114-117.

Allikmets, R., Gerrard, B., Court, D. and Dean, M. (1993). Cloning and organization of the *abc* and *mdl* genes of *Escherichia coli*: relationship to eukaryotic multidrug resistance. Gene *136*, 231-236.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403-410.

Anderson, K.L., Baird, L., Lewis, R.A., Chinault, A.C., Otterud, B., Leppert, M., and Lupski, J.R. (1995). A YAC contig encompassing the recessive Stargardt disease gene (*STGD*) on chromosome 1p. Am. J. Hum. Genet. *57*, 1351-1363.

Anderson, K.L. (July 30, 1996) Towards the Isolation of Genes for Recessively Inherited Ocular Disorders; Bardet-Biedl Syndrome, Leber Congenital Amaurosis, Primary Congenital Glaucoma, and Stargardt Disease. Ph.D. Thesis. Baylor College of Medicine.

Anderson, R.E. and Maude, M.B. (1971). Lipids of ocular tissues: the effects of essential fatty acid deficiency on the phospholipids of the photoreceptor membranes of rat retina. Arch. Biochem. Biophys. *151*, 270-276.

Azarian, S.M., Travis, G.H., (1997) The photoreceptor rim protein is an ABC transporter encoded by the gene for recessive Stargardt's disease (ABCR). FEBS Letters. 409(2):247-52.

Bellanne-Chantelot, C., *et al.* (1992). Mapping the Whole Human Genome by Fingerprinting Yeast Artificial Chromosomes. Cell *70*, 1059-1068.

Birnbach, C.D., Järveläinen, M., Possin, D.E., and Milam, A.H. (1994). Histopathology and immunocytochemistry of the neurosensory retina in fundus flavimaculatus. Ophthalmology *101*, 1211-1219.

Blacharski, D.A. (1988). Fundus flavimaculatus. In Retinal Dystrophies and Degenerations, D.A. Newsome, ed. (New York: Raven Press), pp. 135-159.

Boguski, M.S., Lowe, T.M., and Tolstoshev, C.M. (1993). dbEST-database for "expressed sequence tags". Nature Genet. *4*, 332-333.

Chen, Z.-Y., Battinelli, E.M., Hendricks, R.W., Powell, J.F., Middleton-Price, H., Sims, K.B., Breakfield, X.O., and Craig, I.W. (1993). Norrie disease gene: characterization of deletions and possible function. Genomics *16*, 533-535.

Childs, S., and Ling, V. (1994). The MDR superfamily of genes and its biological implications. In Important Advances in Oncology, V.T. DeVita, S. Hellman and S.A. Rosenberg, eds. (Philadelphia, Pennsylvania: Lippincott Company), pp. 21-36.

Chiu, M. I., Zack, D.J., Wang, Y., and Nathans, J. (1994). Murine and bovine blue cone pigment genes: cloning and characterization of two new members of the S family of visual pigments. Genomics *21*, 440-443.

Chomczynski, P. and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. *162*, 156-159.

Daemen, F.J.M. (1973). Vertebrate rod outer segment membranes. Biochem. Biophys. Acta *300*, 255-288.

de la Salle, H., Hanau, D., Fricker, D., Urlacher, A., Kelly, A., Salamero, J., Powis, S.H., Donato, L., Bausinger, H., Laforet, M., Jeras, M., Spehner, D., Bieber, T., Falkenrodt, A., Cazenave, J.-P., Trowsdale, J., and Tongio, M.-M. (1994). Homozygous human TAP peptide transporter mutation in HLA class I deficiency. Science *265*, 237-241.

Dean, M., and Allikmets, R. (1995). Evolution of ATP-binding cassette transporter genes. Curr. Opin. Genet. Dev. *5*, 779-785.

Dean, M., Allikmets, R., Gerrard, B., Stewart, C., Kistler, A., Shafer, B., Michaelis, S., and Strathern, J. (1994). Mapping and sequencing of two yeast genes belonging to the ATP-binding cassette superfamily. Yeast *10*, 377-383.

Devereaux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. *12*, 387-395.

Dowling, J. E. (1960). Chemistry of visual adaptation in the rat. Nature *188,* 114-118.

Dryja, T.P. and Li, T. (1995). Molecular genetics of retinitis pigmentosa. Hum. Mol. Genet. *4,* 1739-1743.

Eagle, R.C. Jr., Lucier, A.C., Bernadino, V.B. Jr., and Yanoff, M. (1980). Retinal pigment epithelial abnormalities in Fundus Flavimaculatus: A light and electron microscopic study. Ophthalmology *87,* 1189-1200.

Feeney, L. (1978). Lipofuscin and melanin of human retinal pigment epithelium. Fluorescence, enzyme cytochemical, and ultrastructural studies Invest. Ophthalmol. Vis. Sci. *17,* 583-600.

Feng, D.-F., and Doolittle, R.F. (1987). Progressive sequence alignment as a prerequisite to correct phyllogenetic trees. J. Mol. Evol. *25,* 351-360.

Fishman, G.A. (1976). Fundus flavimaculatus: a clinical classification. Arch. Ophthalmol. *94,* 2061-2067.

Fong, S.-L., Liou, G.I., Landers, R.A., Alvarez, R.A., and Bridges, C.D. (1984). Purification and characterization of a retinol-binding glycoprotein synthesized and secreted by bovine neural retina. J. Biol. Chem. *259,* 6534-6542.

Franceschetti, A. (1963). Über tapeto-retinale Degenerationen im Kindesalter. In H. von Sautter, ed. Entwicklung und Fortschritt in der Augenheilkunde. (Stuttgart: Ferdinand Enke) pp. 107-120.

Gass, J.D.M. (1987). Stargardt's disease (Fundus Flavimaculatus) in Stereoscopic Atlas of Macular Diseases: Diagnosis and Treatment, Volume 1, 3rd edition. pages 256-261. The C.V. Mosby Company, St Louis, MO.

Gerber, S., Rozet, J.-M., Bonneau, D., Souied, E., Camuzat, A., Dufier, J.-L., Amalric, P., Weissenbach, J., Munnich, A., and Kaplan, J. (1995). A gene for late-onset fundus flavimaculatus with macular dystrophy maps to chromosome 1p13. Am. J. Hum. Genet. *56*, 396-399.

Glavač, D., and Dean, M. (1993). Optimization of the single-strand conformation polymorphism (SSCP) technique for detection of point mutations. Hum. Mutat. *2*, 404-414.

Hayes, K.C. (1974). Retinal degeneration in monkeys induced by deficiencies of vitamin E or A. Invest. Ophthalmology *13*, 499-510.

Hettema, E.H., van Roermund, C.W.T., Distel, B., van den Berg, M., Vilela, C., Rodrigues-Pousada, C., Wanders, R.J.A., and Tabak, H.F. (1996). The ABC transporter proteins Pat1 and Pat2 are required for import of long-chain fatty acids into peroxisomes of Saccharomyces cerevisiae. EMBO J. *15*, 3813-3822.

Hoyng, C.B., Poppelaars, F., van de Pol, T.J.R., Kremer, H., Pinckers, A.J.L.G., Deutman, A.F., and Cremers, F.P.M. (1996). Genetic fine mapping of the gene for recessive Stargardt disease. Hum. Genet. *98*, 500-504.

Hyde, S.C., Emsley, P., Hartshorn, M.J., Mimmack, M.M., Gileadi, U., Pearce, S.R., Gallagher, M.P., Gill, D.R., Hubbard, R.E., and Higgins, C.F. (1990). Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. Nature *346*, 362-365.

Klein, B.A. and Krill, A.E. (1967). Fundus Flavimaculatus: clinical, functional, and histopathologic observations. Am. J. Ophthalmol. *64*, 3-23.

Klugbauer, N., and Hofmann, F. (1996). Primary structure of a novel ABC transporter with a chromosomal localization on the band encoding the multidrug resistance-associated protein. FEBS Lett. *391*, 61-65.

Kuwano, Y., Nakanishi, O., Nabeshima, Y.-I., Tanaka, T., and Ogata, K. (1985). Molecular cloning and nucleotide sequence of DNA complementary to rat ribosomal protein S26 messenger RNA. J. Biochem. (Tokyo) *97*: 983-992.

Lennon, G., Auffray, C., Polymeropoulos, M. and Soares, M.B. (1996). The I.M.A.G.E. consortium: An integrated molecular analysis of genomes and their expression. Genomics *33*, 151-152.

Lincoln, A.L., Daly, M., and Lander, E. (1991). PRIMER: a computer program for automatically selecting PCR primers. Whitehead Institute Technical Report.

Lopez, P.F., Maumenee, I.H., de la Cruz, Z., Green, W.R. (1990). Autosomal-dominant fundus flavimaculatus: clinicopathologic correlation. Ophthalmology *97*, 798-809.

Luciani, M.-F., Denizot, F., Savary, S., Mattei, M.G., and Chimini, G. (1994). Cloning of two novel ABC transporters mapping on human chromosome 9. Genomics *21*, 150-159.

Luciani, M.-F., and Chimini, G. (1996). The ATP binding cassette transporter ABC1, is required for the engulfment of corpses generated by apoptotic cell death. EMBO J. *15*, 226-235.

McDonnell, P.J., Kivlin, J.D., Maumenee, I.H., and Green, W.R. (1986). Fundus flavimaculatus without maculopathy: a clinicopathologic study. Ophthalmology *93*, 116-119.

Meindl, A., Berger, W., Meitinger, T., van de Pol, D., Achatz, H., Dorner, C., Hasseman, M., Hellebrand, H., Gal, A., Cremers, F., and Ropers, H.H. (1992). Norrie disease is caused by mutations in an extracellular protein resembling c-terminal globular domain of mucins. Nat. Genet. *2*, 139-143.

Meindl, A., Dry, K., Herrmann, K., Manson, F., Ciccodicola, A., Edgar, A., Carvalho, M.R.S., Achatz, H., Hellebrand, H., Lennon, A., Mibliaccio, C., Porter, K., Zrenner, E., Bird, A., Jay,

M., Lorenz, B., Wittwer, B., D'urso, M., Meitinger, T., and Wright, A. (1996). A gene (RPGR) with homology to the RCC1 guanine nucleotide exchange factor is mutated in X-linked retinitis pigmentosa (RP3). Nat. Genet. *13*, 35-42.

Michaelis, S., and Berkower, C. (1995). Sequence comparison of yeast ATP binding cassette (ABC) proteins. In Cold Spring Harbor Symposium on Quantitative Biology, vol.LX: Protein Kinesis - The Dynamics of Protein Trafficking and Stability. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor).

Mosser, J., Douar, A.-M., Sarde, C.-O., Kioschis, P., Feil, R., Moser, H., Poustka, A.-M., Mandel, J.-L., and Aubourg, P. (1993). Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. Nature *361*, 726-730.

Nathans, J., Thomas, D., and Hogness, D.S. (1986). Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. Science *232*, 193-202.

Nickells, R. W., Burgoyne, C.F., Quigley, H.A., and Zack, D.J. (1995). Cloning and characterization of rodopsin cDNA from the old world monkey, Macaca fasciclaris. Investigative Ophthalmology and Visual Science 36:72-82.

Noble, K.G., and Carr, R.E. (1979). Stargardt's disease and fundus flavimaculatus. Arch. Ophthalmol. *97*, 1281-1285.

Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. (1989). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics *5*, 874-879.

Rando, R.R. (1990). The chemistry of vitamin A and vision. Angew. Chem. Int. Ed. Engl. *29*, 461-480.

Riordan, J.R., Rommens, J.M., Kerem, B.-S., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J.-L., Drumm, M.L., Ianuzzi, M.C., Collins, F.S., and Tsui, L.-C. (1989). Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science *245*, 1066-1073.

Roa, B.B., Garcia, C.A., Suter, U., Kulpa, D.A., Wise, C.A., Mueller, J., Welcher, A.A., Snipes, G.J., Shooter, E.M., Patel, P.I., and Lupski, J.R. (1993). Charcot-Marie-Tooth disease type 1A, association with a spontaneous point mutation in the *PMP22* gene. New Eng. J. Med. *329*, 96-101.

Roa, B.B., Warner, L.E., Garcia, C.A., Russo, D., Lovelace, R., Chance P.F., and Lupski, J.R. (1996). Myelin protein zero (*MPZ*) gene mutations in nonduplication type 1 Charcot-Marie-Tooth disease. Hum. Mut. *7*, 36-45.

Rowe, L.B., Nadeau, J.H., Turner, R., Frankel, W.N., Letts, V.A., Eppig, J., Ko, M.S.H., Thurston, S.J., and Birkenmeier, E.H. (1994). Maps from two interspecific backcross DNA panels available as a community genetic mapping resource. Mamm. Genome *5*, 253-274.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Schaeren-Wiemers, N., and Gerfin-Moser, A. (1993). A single protocol to detect transcripts of various types and expression levels in neural tissue and culture cells: in situ hybridization using digoxigenin-labeled cRNA probes. Histochemistry *100*, 431-440.

Seabra, M.C., Brown, M.S., and Goldstein, J.L. (1993). Retinal degeneration in choroideremia: deficiency of Rab geranylgeranyl transferase. Science *259*, 377-381.

Shani, N. and Valle, D. (1996). A *Saccharomyces cerevisiae* homolog of the human adrenoleukodystrophy transporter is a heterodimer of two half ATP-binding cassette transporters. Proc. Natl. Acad. Sci. USA *93*, 11901-11906.

Shimozawa, N., Tsukamoto, T., Suzuki, Y., Orii, T., Shirayoshi, Y., Mori, T., and Fujiki, Y. (1992). A human gene responsible for Zellweger syndrome that affects peroxisome assembly. Science *255*, 1132-1134.

Smit, J.J.M., Schinkel, A.H., Oude Elferink, R.P.J., Groen, A.K., Wagenaar, E., van Deemter, L., Mol, C.A.A.M., Ottenhoff, R., van der Lugt, N.M.T., van Roon, M.A., van der Valk, M.A., Offerhaus, G.J.A., Berns, A.J.M., and Borst, P. (1993). Homozygous disruption of the murine mdr2 P-glycoprotein gene leads to a complete absence of phospholipid from bile and to liver disease. Cell *75*, 451-462.

Stargardt, K. (1909). Über familiäre, progressive Degeneration in der Maculagegend des Auges. Albrecht von Graefes Arch. Klin. Exp. Ophthalmol. *71*, 534-550.

Steinmetz, R.L., Garner, A., Maguire, J.I., and Bird, A.C. (1991). Histopathology of incipient fundus flavimaculatus. Ophthalmology *98*, 953-956.

Stone, E.M., Nichols, B.E., Kimura, A.E., Weingeist, T.A., Drack, A., and Sheffield, V.C. (1994). Clinical features of a Stargardt-like dominant progressive macular dystrophy with genetic linkage to chromosome 6q. Arch. Ophthalmol. *112*, 765-772.

Sun, H., Nathans, J., (1997) Stargardt's ABCR is localized to the disc membrane of retinal rod outer segments. Nature Genetics. 17(1):15-6.

Thomas, P.M., Cote, G.J., Wohllk, N., Haddad, B., Mathew, P.M., Rabl, W., Aguilar-Bryan, L., Gagel, R.F., and Bryan, J. (1995). Mutations in the sulfonylurea receptor gene in familial persistent hyperinsulinemic hypoglycemia of infancy. Science *268*, 426-429.

Valle, D. and Simell, O. (1995). The hyperornithinemias. In The Metabolic and Molecular Basis of Inherited Disease, Scriver, C.R., Beaudet, A.L., Sly, W.S., and Valle, D., eds. (New York: McGraw Hill), pp. 1147-1185.

van Helvoort, A., Smith, A.J., Sprong, H., Fritzsche, I., Schinkel, A.H., Borst, P., and van Meer, G. (1996). MDR1 P-glycoprotein is a lipid translocase of broad specificity, while MDR3 P-glycoprotein specifically translocates phosphatidylcholine. Cell *87*, 507-517.

Wang, Y., Macke, J.P., Abella, B.S., Andreasson, K., Worley, P., Gilbert, D.J., Copeland, N.G., Jenkins, N.A., and Nathans, J. (1996). A large family of putative transmembrane receptors homologous to the product of the Drosophilal tissue polarity gene frizzled. J. Biol. Chem. 271:4468-4476.

Warner, L.E., Hilz, M.J., Appel, S.H., Killian, J.M., Kolodny, E.H., Karpati, G., Carpenter, S., Watters, G.V., Wheeler, C., Witt, D., Bodell, A., Nelis, E., Van Broeckhoven, C., and Lupski, J.R. (1996). Clinical phenotypes of different *MPZ* (P_0) mutations may include Charcot-Marie-Tooth type 1B, Dejerine-Sottas, and congenital hypomyelination. Neuron *17*, 451-460.

Weber, B.H.F., Vogt, G., Pruett, R.C., Stohr, H., and Felbor, U. (1994). Mutations in the tissue inhibitor of metalloproteinase-3 (TIMP3) in patients with Sorsby's fundus dystrophy. Nat. Genet. *8*, 352-356.

Weiter, J.J., Delori, F.C., Wing, G.L., and Fitch, K.A. (1986). Retinal pigment epithelial lipofuscin and melanin and choroidal melanin in human eyes. Invest. Ophthalmol. Vis. Sci. *27*, 145-152.

White, M.B., Carvalho, M., Derse, D., O'Brien, S.J., and Dean, M. (1992). Detecting single base substitutions as heteroduplex polymorphisms. Genomics *12*, 301-306.

Wing, G.L., Blanchard, G.C., and Weiter, J.J. (1978). The topography and age relationship of lipofuscin concentration in the retinal pigment epithelium. Invest. Ophthalmol. Vis. Sci. *17*, 601-607.

Zhang, K., Bither, P.P., Park, R., Donoso, L.A., Seidman, J.G., and Seidman, C.E. (1994). A dominant Stargardt's macular dystrophy locus maps to chromosome 13q34. Arch. Ophthalmol. *112*, 759-764.

Zhou, H., Yoshioka, T., and Nathans, J. (1996). Retina-derived POU-domain factor-1: a complex POU-domain gene implicated in the development of retinal ganglion and amacrine cells. J. Neurosci. *16*, 2261-2274.

Zinn, K. M. and Marmor, M.F. (1979). The Retinal Pigment Epithelium. (Harvard University Press, Cambridge. MA). pp. 521.

What is claimed is:

1. An isolated nucleic acid sequence encoding retina-specific ATP binding cassette transporter.

2. An isolated nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, or a fragment thereof having substantially the same activity.

3. An isolated nucleic acid sequence selected from the group consisting of SEQ ID NOS: 2 or 5, or a fragment thereof having substantially the same activity.

4. An isolated amino acid sequence selected from the group consisting of SEQ ID NO: 3 or 6, or a fragment thereof having substantially the same activity.

5. An isolated amino acid sequence of Figure 3, or a fragment thereof having substantially the same activity.

6. A vector comprising a nucleic acid sequence encoding retina-specific ATP binding cassette transporter.

7. A vector comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, or a fragment thereof having substantially the same activity.

8. A vector comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 2 or 5, or a fragment thereof having substantially the same activity.

9. A vector comprising a nucleic acid sequence encoding an amino acid sequence selected from the group consisting of SEQ ID NOS: 3 or 6.

10. A vector comprising a nucleic acid sequence encoding the amino acid sequence of Figure 3.

11. A host cell capable of expressing a nucleic acid sequence encoding a retina-specific ATP binding cassette transporter.

12. A host cell capable of expressing a nucleic acid sequence of SEQ ID NO: 1.

13. A host cell capable of expressing a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 2 or 5.

14. A host cell capable of expressing a nucleic acid seqeunce encoding an amino acid sequence selected from the group consisting of SEQ ID NOS: 3 or 6.

15. A host cell capable of expressing a nucleic acid seqeunce encoding the amino acid sequence of Figure 3.

16. A cell culture capable of expressing a retina-specific ATP binding cassette transporter.

17. A cell culture capable of expressing a nucleic acid sequence of SEQ ID NO: 1.

18. A cell culture capable of expressing a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 2 or 5.

19. A cell culture capable of expressing a nucleic acid sequence encoding an amino acid sequence selected from the group consisting of SEQ ID NOS: 3 or 6.

20. A cell culture of claim 19 obtained by transforming a cell with an expression vector comprising a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 2 or 5.

21. A cell culture capable of expressing a nucleic acid sequence encoding an amino acid sequence selected from the group consisting of SEQ ID NOS: 3 or 6.

22. A protein preparation comprising an amino acid sequence for retina-specific ATP binding cassette transporter.

23. A protein preparation comprising an amino acid sequence encoded by a sequence of SEQ ID NO: 1.

24. A protein preparation comprising an amino acid sequence encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 2 or 5.

25. A protein preparation comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 3 or 6.

26. A protein preparation comprising an amino acid sequence of Figure 3.

27. A composition comprising an effective amount of a sequence selected from the group consisting of SEQ ID NOS: 2 or 5 or a fragment thereof having substantially similar activity, and a pharmaceutically acceptable carrier.

28. A composition comprising an effective amount of an antisense sequence to a sequence selected from the group consisting of SEQ ID NOS: 2 or 5 or a fragment thereof which fragment has substantially similar activity, and a pharmaceutically acceptable carrier.

29. A method of screening for an agent that alters retina-specific ATP binding cassette transporter comprising combining purified retina-specific ATP binding cassette transporter and at least one agent suspected of altering retina-specific ATP binding cassette transporter and observing an alteration in said purified retina-specific ATP binding cassette transporter.

30. The method of claim 29 wherein said alteration is activation of said purified retina-specific ATP binding cassette transporter observed by a inhibition of a characteristic associated with macular degeneration selected from the group consisting of inhibition of central visual impairment, inhibition of progressive bilateral atrophy of the macular retinal pigment epithelium, inhibition of progressive bilateral atrophy of the neuroepithelium, inhibition of macula flecks, inhibition of midretinal periphery flecks, and inhibition of retina-specific ATP binding cassette transporter transcripts in photoreceptor cells.

31. The method of claim 30 wherein said macular degeneration is selected from the group consisting of Stargardt Disease, Fundus Flavimaculatus, and age-related macular degeneration.

32. A method of claim 29 wherein said alteration is an inhibition of said purified retina-specific ATP binding cassette transporter observed by a characteristic associated with macular degeneration selected from the group consisting of central visual impairment, bilateral atrophy of the macular retinal pigment epithelium, bilateral atrophy of the neuroepithelium, macula flecks, midretinal periphery flecks, and retina-specific ATP binding cassette transporter transcripts in photoreceptor cells.

33. A method of screening for an agent that inhibits macular degeneration comprising combining purified retina-specific ATP binding cassette transporter from a patient suspected of having macular degeneration and at least one agent suspected of activating retina-specific ATP binding cassette transporter and observing an activation in said purified retina-specific ATP binding cassette transporter.

34. A method of screening for an agent that activates macular degeneration comprising combining a purified wild-type retina-specific ATP binding cassette transporter and at least one agent suspected of activating macular degeneration and observing an inhibition in said purified wild-type retina-specific ATP binding cassette transporter.

35. A transgenic non-human mammal comprising a recombinant sequence encoding a retina-specific ATP binding cassette transporter introduced into said mammal, or an ancestor of said mammal.

36. The mammal of claim 35 wherein said sequence encoding said retina-specific ATP binding cassette transporter is selected from the group consisting of SEQ ID NOS: 1, 2, and 5.

37. A transgenic non-human mammal comprising a suppressed retina-specific ATP binding cassette transporter gene.

38. A transgenic non-human mammal comprising a recombinant wild-type sequence encoding retina-specific ATP binding cassette transporter.

39. The transgenic non-human mammal of claim 35 wherein said retina-specific ATP binding cassette transporter sequence is selected from the group consisting of SEQ ID NOS: 3 and 6.

40. A diagnostic kit for detecting macular degeneration comprising in one or more containers a pair of primers, wherein one primer within said pair is complementary to a region of the retina-specific ATP binding cassette receptor, a probe specific to the amplified product, and a means for visualizing amplified DNA, and optionally including one or more size markers, and positive and negative controls.

41. The diagnostic kit of claim 40 wherein said primer is selected from the group consisting of SEQ ID NOS: 12-113.

42. The diagnostic kit of claim 40 wherein said primer is complementary to a region flanking an exon of retina-specific ATP binding cassette receptor genomic DNA sequence.

43. The diagnostic kit of claim 40 wherein said means for visualizing amplified DNA is selected from the group consisting of fluorescent stain, ^{32}P, and biotin.

44. A method of detecting macular degeneration comprising:

obtaining a sample comprising patient nucleic acids from a patient tissue sample;

amplifying retina-specific ATP binding cassette receptor specific nucleic acids from said patient nucleic acids to produce a test fragment;

obtaining a sample comprising control nucleic acids from a control tissue sample;

amplifying control nucleic acids encoding wild-type retina-specific ATP binding cassette receptor to produce a control fragment;

comparing the test fragment with the control fragment to detect the presence of a sequence difference in the test fragment, wherein a difference in said test fragment indicates macular degeneration.

45. The method of claim 44 wherein a sequence difference is selected from the group consisting of a missense mutation, an intragenic deletion, intragenic insertion, a splice donor site mutation, and a frameshift.

46. The method of claim 44 wherein a sequence difference is a missense mutation.

47. The method of claim 44 wherein said amplification step comprises performing the polymerase chain reaction.

48. The method of claim 47 wherein the polymerase chain reaction comprises using a pair of primers, wherein one primer within said pair is selected from the group consisting of SEQ ID NOS: 12-113.

49. The method of claim 44 wherein said tissue sample is selected from the group consisting of blood, skin, serum, saliva, sputum, mucus, bone marrow, urine, lymph, a tear, chorion, and amniotic fluid.

50. The method of claim 44 wherein said sequence difference is selected from the group consisting of 0223T→G, 0634C→T, 0746A→G, 1018T→G, 1411G→A, 1569T→G, 1715G→A, 1715G→C, 1804C→T, 1822T→A, 1917C→A, 2453G→A, 2461T→A, 2536G→C, 2588G→C, 2791G→A, 2827C→T, 2894A→G, 3083C→T, 3212C→T, 3215T→C, 3259G→A, 3322C→T, 3364G→A, 3385G→T, 3386G→T, 3602T→G, 3610G→A, 4139C→T, 4195G→A, 4222T→C, 4297G→A, 4316G→A, 4319T→C, 4346G→A, 4462T→C, 4469G→A, 4577C→T, 4594G→A, 5041del15, 5281del9, 5459G→C, 5512C→T, 5527C→T, 5657G→A, 5693G→A, 5882G→A, 5908C→T, 5929G→A, 6079C→T, 6088C→T, 6089G→A, 6112C→T, 6148G→C, 6166A→T, 6229C→T, 6286G→A, 6316C→T, 6391G→A, 6415C→T, 6445C→T, and 6543del36.

51. The method of claim 44 further wherein said sequence difference results in an amino acid sequence difference selected from the group consisting of C75G, R212C, D249G, Y340D, E471K, D523E, R572Q, R572P, R602W, F608I, Y639X, G818E, W821R, D846H, G863A, V931M, R943W, N965S, A1028V, S1071L,V1072A, E1087K, R1108C, E1122K, R1129C, R1129L, L1201R, D1204N, P1380L, E1399K, W1408R, V1433I, G1439D, F1440S, W1449X, C1488R, C1490Y, T1526M, D1532N, VVAIC1681del, PAL1761del, R1820P, H1838Y, R1843W, G1886E, R1898H, G1961E, L1970F, G1977S, L2027F, R2030X, R2030Q, R2038W, V2050L, K2056X, R2077W, E2096K, R2106C, E2131K, R2139W, R2149X, 1181del12, 0664del13, 2884delC, 4232insTATG, 4947delC, 6709delG, 4253+5G→T, 5196+2T→C, 5585+1G→A, 5714+5G→A, 5898+1G→A, and 6005+1G→T.

52. The method of claim 44 wherein said sequence difference results in a frame shift in the amino acid sequence.

53. The method of claim 44 wherein said sequence difference results in a splice site in the amino acid sequence.

54. A sequence of having a sequence of SEQ ID NOS: 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, or 113.

55. A sequence encoding SEQ ID NO: 2 having a mutation selected from the group consisting of 0223T→G, 0634C→T, 0746A→G, 1018T→G, 1411G→A, 1569T→G, 1715G→A, 1715G→C, 1804C→T, 1822T→A, 1917C→A, 2453G→A, 2461T→A, 2536G→C, 2588G→C, 2791G→A, 2827C→T, 2894A→G, 3083C→T, 3212C→T, 3215T→C, 3259G→A, 3322C→T, 3364G→A, 3385G→T, 3386G→T, 3602T→G, 3610G→A, 4139C→T, 4195G→A, 4222T→C, 4297G→A, 4316G→A, 4319T→C, 4346G→A, 4462T→C, 4469G→A, 4577C→T, 4594G→A, 5041del15, 5281del9, 5459G→C, 5512C→T, 5527C→T, 5657G→A, 5693G→A, 5882G→A, 5908C→T, 5929G→A, 6079C→T, 6088C→T, 6089G→A, 6112C→T, 6148G→C, 6166A→T, 6229C→T, 6286G→A, 6316C→T, 6391G→A, 6415C→T, 6445C→T, and 6543del36.

56. A sequence of claim 55 wherein said sequence difference results in a frame shift in the amino acid sequence.

57. The method of claim 55 wherein said sequence difference results in a splice site in the amino acid sequence.

58. A sequence encoding SEQ ID NO: 3 having a mutation selected from the group consisting of C75G, R212C, D249G, Y340D, E471K, D523E, R572Q, R572P, R602W, F6081, Y639X, G818E, W821R, D846H, G863A, V931M, R943W, N965S, A1028V, S1071L, V1072A, E1087K, R1108C, E1122K, R1129C, R1129L, L1201R, D1204N, P1380L, E1399K, W1408R, V1433I, G1439D, F1440S, W1449X, C1488R, C1490Y, T1526M, D1532N, VVAIC1681del, PAL1761del, R1820P, H1838Y, R1843W, G1886E, R1898H, G1961E, L1970F, G1977S, L2027F, R2030X, R2030Q, R2038W, V2050L, K2056X, R2077W, E2096K, R2106C, E2131K, R2139W, R2149X, 1181del12, 0664del13, 2884delC, 4232insTATG, 4947delC, 6709delG, 4253+5G→T, 5196+2T→C, 5585+1G→A, 5714+5G→A, 5898+1G→A, and 6005+1G→T.

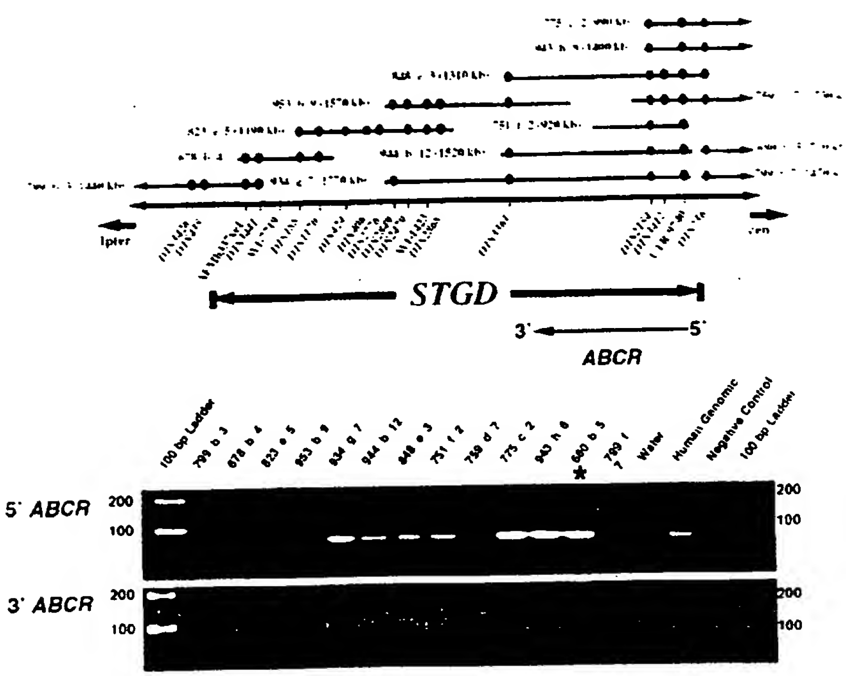

STGD
3'
5'
ABCR
100 bp Ladder
Water
Human Genomic
Negative Control
100 bp Ladder
5' ABCR
3' ABCR
200
100
200
100

Figure 1

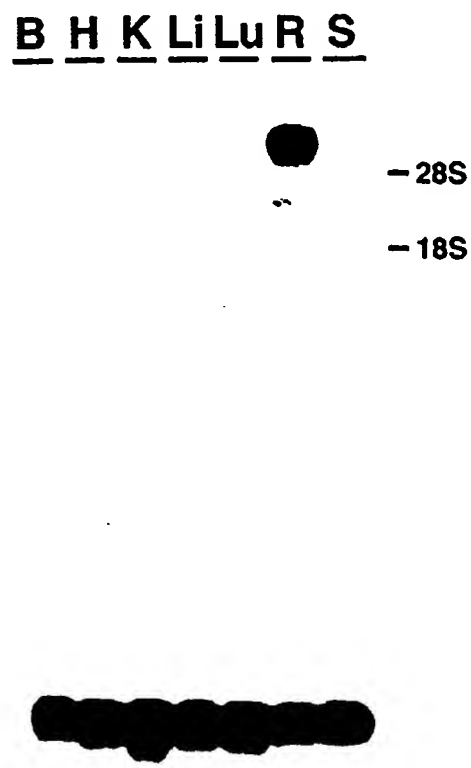


Figure 2

```
     -580                 -560                 -540
CCCCTACCCCTCTGCTAAGCTCAGGGATAACCCAACTAGCTGACCATAATGACTTCAGTC
     -520                 -500                 -480
ATTACGGAGCAAGATGAAAGACTAAAAGAGGGAGGGATCACTTCAGATCTGCCGAGTGAG
     -460                 -440                 -420
TCGATTGGACTTAAAGGGCCAGTCAAACCCTGACTGCCGGCTCATGGCAGGCTCTTGCCG
     -400                 -380                 -360
AGGACAAATGCCCAGCCTATATTTATGCAAAGAGATTTTGTTCCAAACTTAAGGTCAAAG
     -340                 -320                 -300
ATACCTAAAGACATCCCCCTCAGGAACCCCTCTCATGGAGGAGAGTGCCTGAGGGTCTTG
     -280                 -260                 -240
GTTTCCCATTGCATCCCCCACCTCAATTTCCCTGGTGCCCAGCCACTTGTGTCTTTAGGG
     -220                 -200                 -180
TTCTCTTTCTCTCCATAAAAGGGAGCCAACACAGTGTCGGCCTCCTCTCCCCAACTAAGG
     -160                 -140                 -120
GCTTATGTGTAATTAAAAGGGATTATGCTTTGAAGGGGAAAAGTAGCCTTTAATCACCAG
     -100                 -80                  -60
GAGAAGGACACAGCGTCCGGAGCCAGAGGCGCTCTTAACGGCGTTTATGTCCTTTGCTGT
     -40                  -20                  0
CCTGAGGGGCCTCAGCTCTGACCAATCTGGTCTTCGTGTGGTCATTAGCATGGGCTTCGT
                                                M  G  F  V
     20                   40                   60
GAGACAGATACAGCTTTTGCTCTGGAAGAACTGGACCCTGCGGAAAAGGCAAAAG|ATTCG
 R  Q  I  Q  L  L  L  W  K  N  W  T  L  R  K  R  Q  K   I  R
     80                   100                  120
CTTTGTGGTGGAACTCGTGTGGCCTTTATCTTTATTTCTGGTCTTGATCTGGTTAAGGAA
 F  V  V  E  L  V  W  P  L  S  L  F  L  V  L  I  W  L  R  N
     140                  160                  180
TGCCAACCCGCTCTACAGCCATCATGAAT|GCCATTTCCCCAACAAGGCGATGCCCTCAGC
 A  N  P  L  Y  S  H  H  E  C   H  F  P  N  K  A  M  P  S  A
     200                  220                  240
AGGAATGCTGCCGTGGCTCCAGGGGATCTTCTGCAATGTGAACAATCCCTGTTTTCAAAG
 G  M  L  P  W  L  Q  G  I  F  C  N  V  N  N  P  C  F  Q  S
     260                  280                  300
CCCCACCCCAGGAGAATCTCCTGGAATTGTGTCAAACTATAACAACTCCAT|CTTGGCAAG
 P  T  P  G  E  S  P  G  I  V  S  N  Y  N  N  S  I   L  A  R
     320                  340                  360
GGTATATCGAGATTTTCAAGAACTCCTCATGAATGCACCAGAGAGCCAGCACCTTGGCCG
 V  Y  R  D  F  Q  E  L  L  M  N  A  P  E  S  Q  H  L  G  R
     380                  400                  420
TATTTGGACAGAGCTACACATCTTGTCCCAATTCATGGACACCCTCCGGACTCACCCGGA
 I  W  T  E  L  H  I  L  S  Q  F  M  D  T  L  R  T  H  P  E
     440                  460                  480
GAGAATTGCAG|GAAGAGGAATACGAATAAGGGATATCTTGAAAGATGAAGAAACACTGAC
 R  I  A  G   R  G  I  R  I  R  D  I  L  K  D  E  E  T  L  T
     500                  520                  540
ACTATTTCTCATTAAAAACATCGGCCTGTCTGACTCAGTGGTCTACCTTCTGATCAACTC
 L  F  L  I  K  N  I  G  L  S  D  S  V  V  Y  L  L  I  N  S
```

FIGURE 3A

```
      560                 580                 600
TCAAGTCCGTCCAGAGCAG|TTCGCTCATGGAGTCCCGGACCTGGCGCTGAAGGACATCGC
 Q  V  R  P  E  Q   F  A  H  G  V  P  D  L  A  L  K  D  I  A
      620                 640                 660
CTGCAGCGAGGCCCTCCTGGAGCGCTTCATCATCTTCAGCCAGAGACGCGGGGCAAAGAC
 C  S  E  A  L  L  E  R  F  I  I  F  S  Q  R  R  G  A  K  T
      680                 700                 720
GGTGCGCTATGCCCTGTGCTCCCTCTCCCAGGGCACCCTACAGTGGATAGAAGACACTCT
 V  R  Y  A  L  C  S  L  S  Q  G  T  L  Q  W  I  E  D  T  L
      740                 760                 780
GTATGCCAACGTGGACTTCTTCAAGCTCTTCCGTGTG|CTTCCCACACTCCTAGACAGCCG
 Y  A  N  V  D  F  F  K  L  F  R  V   L  P  T  L  L  D  S  R
      800                 820                 840
TTCTCAAGGTATCAATCTGAGATCTTGGGGAGGAATATTATCTGATATGTCACCAAGAAT
 S  Q  G  I  N  L  R  S  W  G  G  I  L  S  D  M  S  P  R  I
      860                 880                 900
TCAAGAG|TTTATCCATCGGCCGAGTATGCAGGACTTGCTGTGGGTGACCAGGCCCCTCAT
 Q  E   F  I  H  R  P  S  M  Q  D  L  L  W  V  T  R  P  L  M
      920                 940                 960
GCAGAATGGTGGTCCAGAGACCTTTACAAAGCTGATGGGCATCCTGTCTGACCTCCTGTG
 Q  N  G  G  P  E  T  F  T  K  L  M  G  I  L  S  D  L  L  C
      980                1000                1020
TGGCTACCCCGAGGGAGGTGGCTCTCGGGTGCTCTCCTTCAACTGGTATGAAGACAATAA
 G  Y  P  E  G  G  G  S  R  V  L  S  F  N  W  Y  E  D  N  N
      1040               1060                1080
CTATAAGGCCTTTCTGGGGATTGACTCCACAAGGAAGGATCCTATCTATTCTTATGACAG
 Y  K  A  F  L  G  I  D  S  T  R  K  D  P  I  Y  S  Y  D  R
      1100               1120                1140
AAGAACAA|CATCCTTTTGTAATGCATTGATCCAGAGCCTGGAGTCAAATCCTTTAACCAA
 R  T  T   S  F  C  N  A  L  I  Q  S  L  E  S  N  P  L  T  K
      1160               1180                1200
AATCGCTTGGAGGGCGGCAAAGCCTTTGCTGATGGGAAAAATCCTGTACACTCCTGATTC
 I  A  W  R  A  A  K  P  L  L  M  G  K  I  L  Y  T  P  D  S
      1220               1240                1260
ACCTGCAGCACGAAGGATACTGAAGAAT|GCCAACTCAACTTTTGAAGAACTGGAACACGT
 P  A  A  R  R  I  L  K  N   A  N  S  T  F  E  E  L  E  H  V
      1280               1300                1320
TAGGAAGTTGGTCAAAGCCTGGGAAGAAGTAGGGCCCCAGATCTGGTACTTCTTTGACAA
 R  K  L  V  K  A  W  E  E  V  G  P  Q  I  W  Y  F  F  D  N
      1340               1360                1380
CAGCACACAGATGAACATGATCAGA|GATACCCTGGGGAACCCAACAGTAAAAGACTTTTT
 S  T  Q  M  N  M  I  R   D  T  L  G  N  P  T  V  K  D  F  L
      1400               1420                1440
GAATAGGCAGCTTGGTGAAGAAGGTATTACTGCTGAAGCCATCCTAAACTTCCTCTACAA
 N  R  Q  L  G  E  E  G  I  T  A  E  A  I  L  N  F  L  Y  K
      1460               1480                1500
GGGCCCTCGGGAAAGCCAGGCTGACGACATGGCCAACTTCGACTGGAGGGACATATTTAA
 G  P  R  E  S  Q  A  D  D  M  A  N  F  D  W  R  D  I  F  N
      1520               1540                1560
CATCACTGATCGCACCCTCCGCCTGGTCAATCAATACCTGGAG|TGCTTGGTCCTGGATAA
 I  T  D  R  T  L  R  L  V  N  Q  Y  L  E   C  L  V  L  D  K
```

FIGURE 3B

```
     1580                 1600                1620
GTTTGAAAGCTACAATGATGAAACTCAGCTCACCCAACGTGCCCTCTCTCTACTGGAGGA
 F  E  S  Y  N  D  E  T  Q  L  T  Q  R  A  L  S  L  L  E  E
     1640                 1660                1680
AAACATGTTCTGGGCCGGAGTGGTATTCCCTGACATGTATCCCTGGACCAGCTCTCTACC
 N  M  F  W  A  G  V  V  F  P  D  M  Y  P  W  T  S  S  L  P
     1700                 1720                1740
ACCCCACGTGAAGTATAAGATCCGAATGGACATAGACGTGGTGGAGAAAACCAATAAGAT
 P  H  V  K  Y  K  I  R  M  D  I  D  V  V  E  K  T  N  K  I
     1760                 1780                1800
TAAAGACAG|GTATTGGGATTCTGGTCCCAGAGCTGATCCCGTGGAAGATTTCCGGTACAT
 K  D  R   Y  W  D  S  G  P  R  A  D  P  V  E  D  F  R  Y  I
     1820                 1840                1860
CTGGGGCGGGTTTGCCTATCTGCAGGACATGGTTGAACAGGGGATCACAAGGAGCCAGGT
 W  G  G  F  A  Y  L  Q  D  M  V  E  Q  G  I  T  R  S  Q  V
     1880                 1900                1920
GCAGGCGGAGGCTCCAGTTGGAATCTACCTCCAGCAGATGCCCTACCCCTGCTTCGTGGA
 Q  A  E  A  P  V  G  I  Y  L  Q  Q  M  P  Y  P  C  F  V  D
     1940                 1960                1980
CGATTC|TTTCATGATCATCCTGAACCGCTGTTTCCCTATCTTCATGGTGCTGGCATGGAT.
 D  S   F  M  I  I  L  N  R  C  F  P  I  F  M  V  L  A  W  I
     2000                 2020                2040
CTACTCTGTCTCCATGACTGTGAAGAGCATCGTCTTGGAGAAGGAGTTGCGACTGAAGGA
 Y  S  V  S  M  T  V  K  S  I  V  L  E  K  E  L  R  L  K  E
     2060                 2080                2100
GACCTTGAAAAATCAGGGTGTCTCCAATGCAGTGATTTGGTGTACCTGGTTCCTGGACAG
 T  L  K  N  Q  G  V  S  N  A  V  I  W  C  T  W  F  L  D  S
     2120                 2140                2160
CTTCTCCATCATGTCGATGAGCATCTTCCTCCTGACGATATTCATCATG|CATGGAAGAAT
 F  S  I  M  S  M  S  I  F  L  L  T  I  F  I  M   H   G  R  I
     2180                 2200                2220
CCTACATTACAGCGACCCATTCATCCTCTTCCTGTTCTTGTTGGCTTTCTCCACTGCCAC
 L  H  Y  S  D  P  F  I  L  F  L  F  L  L  A  F  S  T  A  T
     2240                 2260                2280
CATCATGCTGTGCTTTCTGCTCAGCACCTTCTTCTCCAAGGCCAGTCTGGCAGCAGCCTG
 I  M  L  C  F  L  L  S  T  F  F  S  K  A  S  L  A  A  A  C
     2300                 2320                2340
TAGTGGTGTCATCTATTTCACCCTCTACCTGCCACACATCCTGTGCTTCGCCTGGCAGGA
 S  G  V  I  Y  F  T  L  Y  L  P  H  I  L  C  F  A  W  Q  D
     2360                 2380                2400
CCGCATGACCGCTGAGCTGAAGAAGGCTGTG|AGCTTACTGTCTCCGGTGGCATTTGGATT
 R  M  T  A  E  L  K  K  A  V   S   L  L  S  P  V  A  F  G  F
     2420                 2440                2460
TGGCACTGAGTACCTGGTTCGCTTTGAAGAGCAAGGCCTGGGGCTGCAGTGGAGCAACAT
 G  T  E  Y  L  V  R  F  E  E  Q  G  L  G  L  Q  W  S  N  I
     2480                 2500                2520
CGGGAACAGTCCCACGGAAGGGGACGAATTCAGCTTCCTGCTGTCCATGCAGATGATGCT
 G  N  S  P  T  E  G  D  E  F  S  F  L  L  S  M  Q  M  M  L
     2540                 2560                2580
CCTTGATGCTGCGTGCTATGGCTTACTCGCTTGGTACCTTGATCAGGTGTTTCCAG|GAGA
 L  D  A  A  C  Y  G  L  L  A  W  Y  L  D  Q  V  F  P  G   D
```

FIGURE 3C

```
    2600                    2620                    2640
CTATGGAACCCCACTTCCTTGGTACTTTCTTCTACAAGAGTCGTATTGGCTTGGCGGTGA
 Y  G  T  P  L  P  W  Y  F  L  L  Q  E  S  Y  W  L  G  G  E
    2660                    2680                    2700
AG|GGTGTTCAACCAGAGAAGAAAGAGCCCTGGAAAAGACCGAGCCCCTAACAGAGGAAAC
 G   C  S  T  R  E  E  R  A  L  E  K  T  E  P  L  T  E  E  T
    2720                    2740                    2760
GGAGGATCCAGAGCACCCAGAAGGAATACACG|ACTCCTTCTTTGAACGTGAGCATCCAGG
 E  D  P  E  H  P  E  G  I  H  D   S  F  F  E  R  E  H  P  G
    2780                    2800                    2820
GTGGGTTCCTGGGGTATGCGTGAAGAATCTGGTAAAGATTTTTGAGCCCTGTGGCCGGCC
 W  V  P  G  V  C  V  K  N  L  V  K  I  F  E  P  C  G  R  P
    2840                    2860                    2880
AGCTGTGGACCGTCTGAACATCACCTTCTACGAGAACCAGATCACCGCATTCCTGGGCCA
 A  V  D  R  L  N  I  T  F  Y  E  N  Q  I  T  A  F  L  G  H
    2900                    2920                    2940
CAATGGAGCTGGGAAAACCACCACCTT|GTCCATCCTGACGGGTCTGTTGCCACCAACCTC
 N  G  A  G  K  T  T  T  L   S  I  L  T  G  L  L  P  P  T  S
    2960                    2980                    3000
TGGGACTGTGCTCGTTGGGGGAAGGGACATTGAAACCAGCCTGGATGCAGTCCGGCAGAG
 G  T  V  L  V  G  G  R  D  I  E  T  S  L  D  A  V  R  Q  S
    3020                    3040                    3060
CCTTGGCATGTGTCCACAGCACAACATCCTGTTCCACCA|CCTCACGGTGGCTGAGCACAT
 L  G  M  C  P  Q  H  N  I  L  F  H  H   L  T  V  A  E  H  M
    3080                    3100                    3120
GCTGTTCTATGCCCAGCTGAAAGGAAAGTCCCAGGAGGAGGCCCAGCTGGAGATGGAAGC
 L  F  Y  A  Q  L  K  G  K  S  Q  E  E  A  Q  L  E  M  E  A
    3140                    3160                    3180
CATGTTGGAGGACACAGGCCTCCACCACAAGCGGAATGAAGAGGCTCAGGACCTATCAG|G
 M  L  E  D  T  G  L  H  H  K  R  N  E  E  A  Q  D  L  S  G
    3200                    3220                    3240
TGGCATGCAGAGAAAGCTGTCGGTTGCCATTGCCTTTGTGGGAGATGCCAAGGTGGTGAT
 G  M  Q  R  K  L  S  V  A  I  A  F  V  G  D  A  K  V  V  I
    3260                    3280                    3300
TCTGGACGAACCCACCTCTGGGGTGGACCCTTACTCGAGACGCTCAATCTGGGATCTGCT
 L  D  E  P  T  S  G  V  D  P  Y  S  R  R  S  I  W  D  L  L
    3320                    3340                    3360
CCTGAAGTATCGCTCAG|GCAGAACCATCATCATGTCCACTCACCACATGGACGAGGCCGA
 L  K  Y  R  S  G   R  T  I  I  M  S  T  H  H  M  D  E  A  D
    3380                    3400                    3420
CCTCCTTGGGGACCGCATTGCCATCATTGCCCAGGGAAGGCTCTACTGCTCAGGCACCCC
 L  L  G  D  R  I  A  I  I  A  Q  G  R  L  Y  C  S  G  T  P
    3440                    3460                    3480
ACTCTTCCTGAAGAACTGCTTTGGCACAGGCTTGTACTTAACCTTGGTGCGCAAGATGAA
 L  F  L  K  N  C  F  G  T  G  L  Y  L  T  L  V  R  K  M  K
    3500                    3520                    3540
AAACATCCAGAGCCAAAGGAAAGGCAGTGAG|GGGACCTGCAGCTGCTCGTCTAAGGGTTT
 N  I  Q  S  Q  R  K  G  S  E   G  T  C  S  C  S  S  K  G  F
    3560                    3580                    3600
CTCCACCACGTGTCCAGCCCACGTCGATGACCTAACTCCAGAACAAGTCCTGGATG|GGGA
 S  T  T  C  P  A  H  V  D  D  L  T  P  E  Q  V  L  D  G   D
```

FIGURE 3D

```
     3620                     3640                     3660
TGTAAATGAGCTGATGGATGTAGTTCTCCACCATGTTCCAGAGGCAAAGCTGGTGGAGTG
 V  N  E  L  M  D  V  V  L  H  H  V  P  E  A  K  L  V  E  C
     3680                     3700                     3720
CATTGGTCAAGAACTTATCTTCCTTCTTCCAAATAAGAACTTCAAGCACAGAGCATATGC
 I  G  Q  E  L  I  F  L  L  P  N  K  N  F  K  H  R  A  Y  A
     3740                     3760                     3780
CAGCCTTTTCAGAGAGCTGGAGGAGACGCTGGCTGACCTTGGTCTCAGCAGTTTTGGAAT
 S  L  F  R  E  L  E  E  T  L  A  D  L  G  L  S  S  F  G  I
     3800                     3820                     3840
TTCTGACACTCCCCTGGAAGAG|ATTTTTCTGAAGGTCACGGAGGATTCTGATTCAGGACC
 S  D  T  P  L  E  E   I  F  L  K  V  T  E  D  S  D  S  G  P
     3860                     3880                     3900
TCTGTTTGCGG|GTGGCGCTCAGCAGAAAAGAGAAAACGTCAACCCCCGACACCCCTGCTT
 L  F  A  G   G  A  Q  Q  K  R  E  N  V  N  P  R  H  P  C  L
     3920                     3940                     3960
GGGTCCCAGAGAGAAGGCTGGACAGACACCCCAGGACTCCAATGTCTGCTCCCCAGGGGC
 G  P  R  E  K  A  G  Q  T  P  Q  D  S  N  V  C  S  P  G  A
     3980                     4000                     4020
GCCGGCTGCTCACCCAGAGGGCCAGCCTCCCCCAGAGCCAGAGTGCCCAGGCCCGCAGCT
 P  A  A  H  P  E  G  Q  P  P  P  E  P  E  C  P  G  P  Q  L
     4040                     4060                     4080
CAACACGGGGACACAGCTGGTCCTCCAGCATGTGCAGGCGCTGCTGGTCAAGAGATTCCA
 N  T  G  T  Q  L  V  L  Q  H  V  Q  A  L  L  V  K  R  F  Q
     4100                     4120                     4140
ACACACCATCCGCAGCCACAAGGACTTCCTGGCGCAG|ATCGTGCTCCCGGCTACCTTTGT
 H  T  I  R  S  H  K  D  F  L  A  Q   I  V  L  P  A  T  F  V
     4160                     4180                     4200
GTTTTTGGCTCTGATGCTTTCTATTGTTATCCTTCCTTTTGGCGAATACCCCGCTTTGAC
 F  L  A  L  M  L  S  I  V  I  L  P  F  G  E  Y  P  A  L  T
     4220                     4240                     4260
CCTTCACCCCTGGATATATGGGCAGCAGTACACCTTCTTCAG|CATGGATGAACCAGGCAG
 L  H  P  W  I  Y  G  Q  Q  Y  T  F  F  S   M  D  E  P  G  S
     4280                     4300                     4320
TGAGCAGTTCACGGTACTTGCAGACGTCCTCCTGAATAAGCCAGGCTTTGGCAACCGCTG
 E  Q  F  T  V  L  A  D  V  L  L  N  K  P  G  F  G  N  R  C
     4340                     4360                     4380
CCTGAAGGAAGGGTGGCTTCC|GGAGTACCCCTGTGGCAACTCAACACCCTGGAAGACTCC
 L  K  E  G  W  L  P   E  Y  P  C  G  N  S  T  P  W  K  T  P
     4400                     4420                     4440
TTCTGTGTCCCCAAACATCACCCAGCTGTTCCAGAAGCAGAAATGGACACAGGTCAACCC
 S  V  S  P  N  I  T  Q  L  F  Q  K  Q  K  W  T  Q  V  N  P
     4460                     4480                     4500
TTCACCATCCTGCAG|GTGCAGCACCAGGGAGAAGCTCACCATGCTGCCAGAGTGCCCCGA
 S  P  S  C  R   C  S  T  R  E  K  L  T  M  L  P  E  C  P  E
     4520                     4540                     4560
GGGTGCCGGGGGCCTCCCGCCCCCCCAG|AGAACACAGCGCAGCACGGAAATTCTACAAGA
 G  A  G  G  L  P  P  P  Q   R  T  Q  R  S  T  E  I  L  Q  D
```

FIGURE 3E

```
    4580                    4600                    4620
CCTGACGGACAGGAACATCTCCGACTTCTTGGTAAAAACGTATCCTGCTCTTATAAGAAG
 L  T  D  R  N  I  S  D  F  L  V  K  T  Y  P  A  L  I  R  S

    4640                    4660                    4680
CAG|CTTAAAGAGCAAATTCTGGGTCAATGAACAGAG|GTATGGAGGAATTTCCATTGGAGG
 S  L  K  S  K  F  W  V  N  E  Q  R   Y  G  G  I  S  I  G  G
    4700                    4720                    4740
AAAGCTCCCAGTCGTCCCCATCACGGGGGAAGCACTTGTTGGGTTTTTAAGCGACCTTGG
 K  L  P  V  V  P  I  T  G  E  A  L  V  G  F  L  S  D  L  G
    4760                    4780                    4800
CCGGATCATGAATGTGAGCGGG|GGCCCTATCACTAGAGAGGCCTCTAAAGAAATACCTGA
 R  I  M  N  V  S  G   G  P  I  T  R  E  A  S  K  E  I  P  D
    4820                    4840                    4860
TTTCCTTAAACATCTAGAAACTGAAGACAACATTA|AGGTGTGGTTTAATAACAAAGGCTG
 F  L  K  H  L  E  T  E  D  N  I  K   V  W  F  N  N  K  G  W
    4880                    4900                    4920
GCATGCCCTGGTCAGCTTTCTCAATGTGGCCCACAACGCCATCTTACGGGCCAGCCTGCC
 H  A  L  V  S  F  L  N  V  A  H  N  A  I  L  R  A  S  L  P
    4940                    4960                    4980
TAAGGACAGGAGCCCCGAGGAGTATGGAATCACCGTCATTAGCCAACCCCTGAACCTGAC
 K  D  R  S  P  E  E  Y  G  I  T  V  I  S  Q  P  L  N  L  T
    5000                    5020                    5040
CAAGGAGCAGCTCTCAGAGATTACAGT|GCTGACCACTTCAGTGGATGCTGTGGTTGCCAT
 K  E  Q  L  S  E  I  T  V   L  T  T  S  V  D  A  V  V  A  I
    5060                    5080                    5100
CTGCGTGATTTTCTCCATGTCCTTCGTCCCAGCCAGCTTTGTCCTTTATTTGATCCAGGA
 C  V  I  F  S  M  S  F  V  P  A  S  F  V  L  Y  L  I  Q  E
    5120                    5140                    5160
GCGGGTGAACAAATCCAAGCACCTCCAGTTTATCAGTGGAGTGAGCCCCACCACCTACTG
 R  V  N  K  S  K  H  L  Q  F  I  S  G  V  S  P  T  T  Y  W
    5180                    5200                    5220
GGTGACCAACTTCCTCTGGGACATC|ATGAATTATTCCGTGAGTGCTGGGCTGGTGGTGGG
 V  T  N  F  L  W  D  I   M  N  Y  S  V  S  A  G  L  V  V  G
    5240                    5260                    5280
CATCTTCATCGGGTTTCAGAAGAAAGCCTACACTTCTCCAGAAAACCTTCCTGCCCTTGT
 I  F  I  G  F  Q  K  K  A  Y  T  S  P  E  N  L  P  A  L  V
    5300                    5320                    5340
GGCACTGCTCCTGCTGTATGG|ATGGGCGGTCATTCCCATGATGTACCCAGCATCCTTCCT
 A  L  L  L  L  Y  G   W  A  V  I  P  M  M  Y  P  A  S  F  L
    5360                    5380                    5400
GTTTGATGTCCCCAGCACAGCCTATGTGGCTTTATCTTGTGCTAATCTGTTCATCGGCAT
 F  D  V  P  S  T  A  Y  V  A  L  S  C  A  N  L  F  I  G  I
    5420                    5440                    5460
CAACAGCAGTGCTATTACCTTCATCTTGGAATTATTTGATAATAACCGG|ACGCTGCTCAG
 N  S  S  A  I  T  F  I  L  E  L  F  D  N  N  R   T  L  L  R
    5480                    5500                    5520
GTTCAACGCCGTGCTGAGGAAGCTGCTCATTGTCTTCCCCCACTTCTGCCTGGGCCGGGG
 F  N  A  V  L  R  K  L  L  I  V  F  P  H  F  C  L  G  R  G
```

FIGURE 3F

```
        5540                    5560                    5580
CCTCATTGACCTTGCACTGAGCCAGGCTGTGACAGATGTCTATGCCCGGTTTG|GTGAGGA
 L  I  D  L  A  L  S  Q  A  V  T  D  V  Y  A  R  F  G   E  E
        5600                    5620                    5640
GCACTCTGCAAATCCGTTCCACTGGGACCTGATTGGGAAGAACCTGTTTGCCATGGTGGT
 H  S  A  N  P  F  H  W  D  L  I  G  K  N  L  F  A  M  V  V
        5660                    5680                    5700
GGAAGGGGTGGTGTACTTCCTCCTGACCCTGCTGGTCCAGCGCCACTTCTTCCTCTCCCA
 E  G  V  V  Y  F  L  L  T  L  L  V  Q  R  H  F  F  L  S  Q
        5720                    5740                    5760
ATG|GATTGCCGAGCCCACTAAGGAGCCCATTGTTGATGAAGATGATGATGTGGCTGAAGA
 W   I  A  E  P  T  K  E  P  I  V  D  E  D  D  D  V  A  E  E
        5780                    5800                    5820
AAGACAAAGAATTATTACTGGTGGAAATAAAACTGACATCTTAAGGCTACATGAACTAAC
 R  Q  R  I  I  T  G  G  N  K  T  D  I  L  R  L  H  E  L  T
        5840                    5860                    5880
CAAG|ATTTATCTGGGCACCTCCAGCCCAGCAGTGGACAGGCTGTGTGTCGGAGTTCGCCC
 K   I  Y  L  G  T  S  S  P  A  V  D  R  L  C  V  G  V  R  P
        5900                    5920                    5940
TGGAGAG|TGCTTTGGCCTCCTGGGAGTGAATGGTGCCGGCAAAACAACCACATTCAAGAT
 G  E   C  F  G  L  L  G  V  N  G  A  G  K  T  T  T  F  K  M
        5960                    5980                    6000
GCTCACTGGGGACACCACAGTGACCTCAGGGGATGCCACCGTAGCAGGCAAGAG|TATTTT
 L  T  G  D  T  T  V  T  S  G  D  A  T  V  A  G  K  S   I  L
        6020                    6040                    6060
AACCAATATTTCTGAAGTCCATCAAAATATGGGCTACTGTCCTCAGTTTGATGCAATCGA
 T  N  I  S  E  V  H  Q  N  M  G  Y  C  P  Q  F  D  A  I  D
        6080                    6100                    6120
TGAGCTGCTCACAGGACGAGAACATCTTTACCTTTATGCCCGGCTTCGAGGTGTACCAGC
 E  L  L  T  G  R  E  H  L  Y  L  Y  A  R  L  R  G  V  P  A
        6140                    6160                    6180
AGAAGAAATCGAAAAG|GTTGCAAACTGGAGTATTAAGAGCCTGGGCCTGACTGTCTACGC
 E  E  I  E  K   V  A  N  W  S  I  K  S  L  G  L  T  V  Y  A
        6200                    6220                    6240
CGACTGCCTGGCTGGCACGTACAGTGGGGGCAACAAGCGGAAACTCTCCACAGCCATCGC
 D  C  L  A  G  T  Y  S  G  G  N  K  R  K  L  S  T  A  I  A
        6260                    6280                    6300
ACTCATTGGCTGCCCACCGCTGGTGCTGCTG|GATGAGCCCACCACAGGGATGGACCCCCA
 L  I  G  C  P  P  L  V  L  L   D  E  P  T  T  G  M  D  P  Q
        6320                    6340                    6360
GGCACGCCGCATGCTGTGGAACGTCATCGTGAGCATCATCAGAGAAGGGAGGGCTGTGGT
 A  R  R  M  L  W  N  V  I  V  S  I  I  R  E  G  R  A  V  V
        6380                    6400                    6420
CCTCACATCCCACAG|CATGGAAGAATGTGAGGCACTGTGTACCCGGCTGGCCATCATGGT
 L  T  S  H  S   M  E  E  C  E  A  L  C  T  R  L  A  I  M  V
        6440                    6460                    6480
AAAGGGCGCCTTTCGATGTATGGGCACCATTCAGCATCTCAAGTCCAA|ATTTGGAGATGG
 K  G  A  F  R  C  M  G  T  I  Q  H  L  K  S  K   F  G  D  G
        6500                    6520                    6540
CTATATCGTCACAATGAAGATCAAATCCCCGAAGGACGACCTGCTTCCTGACCTGAACCC
 Y  I  V  T  M  K  I  K  S  P  K  D  D  L  L  P  D  L  N  P
```

FIGURE 3G

```
     6560                6580                6600
TGTGGAGCAGTTCTTCCAGGGGAACTTCCCAGGCAGTGTGCAGAGGGAGAGGCACTACAA
 V  E  Q  F  F  Q  G  N  F  P  G  S  V  Q  R  E  R  H  Y  N
     6620                6640                6660
CATGCTCCAGTTCCAGGTCTCCTCCTCCCTGGCGAGGATCTTCCAGCTCCTCCTCTC
 M  L  Q  F  Q  V  S  S  S  S  L  A  R  I  F  Q  L  L  L  S
     6680                6700                6720
CCACAAGGACAGCCTGCTCATCGAGGAGTACTCAGTCACACAGACCACACTGGACCAG|GT
 H  K  D  S  L  L  I  E  E  Y  S  V  T  Q  T  T  L  D  Q  V
     6740                6760                6780
GTTTGTAAATTTTGCTAAACAGCAGACTGAAAGTCATGACCTCCCTCTGCACCCTCGAGC
 F  V  N  F  A  K  Q  Q  T  E  S  H  D  L  P  L  H  P  R  A
     6800                6820                6840
TGCTGGAGCCAGTCGACAAGCCCAG|GACTGATCTTTCACACCGCTCGTTCCTGCAGCCAG
 A  G  A  S  R  Q  A  Q  D  *
     6860                6880                6900
AAAGGAACTCTGGGCAGCTGGAGGCGCAGGAGCCTGTGCCCATATGGTCATCCAAATGGA
     6920                6940                6960
CTGGCCCAGCGTAAATGACCCCACTGCAGCAGAAAACAAACACACGAGGAGCATGCAGCG
     6980                7000                7020
AATTCAGAAAGAGGTCTTTCAGAAGGAAACCGAAACTGACTTGCTCACCTGGAACACCTG
     7040                7060                7080
ATGGTGAAACCAAACAAATACAAAATCCTTCTCCAGACCCCAGAACTAGAAACCCCGGGC
     7100                7120                7140
CATCCCACTAGCAGCTTTGGCCTCCATATTGCTCTCATTTCAAGCAGATCTGCTTTTCTG
     7160                7180
CATGTTTGTCTGTGTGTCTGCGTTGTGTGTGATTTTCATGGAAA
```

FIGURE 3H

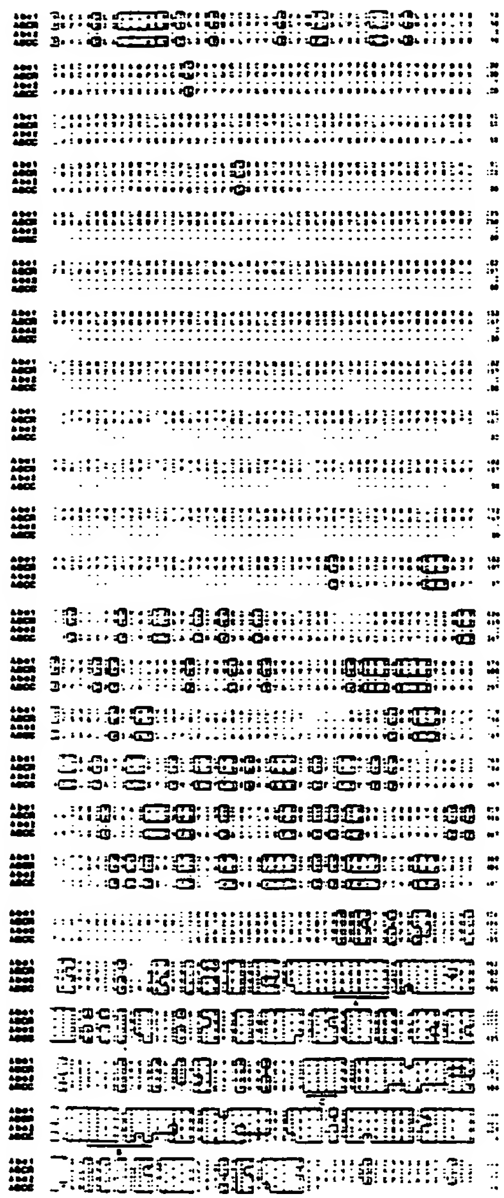


FIGURE 4A

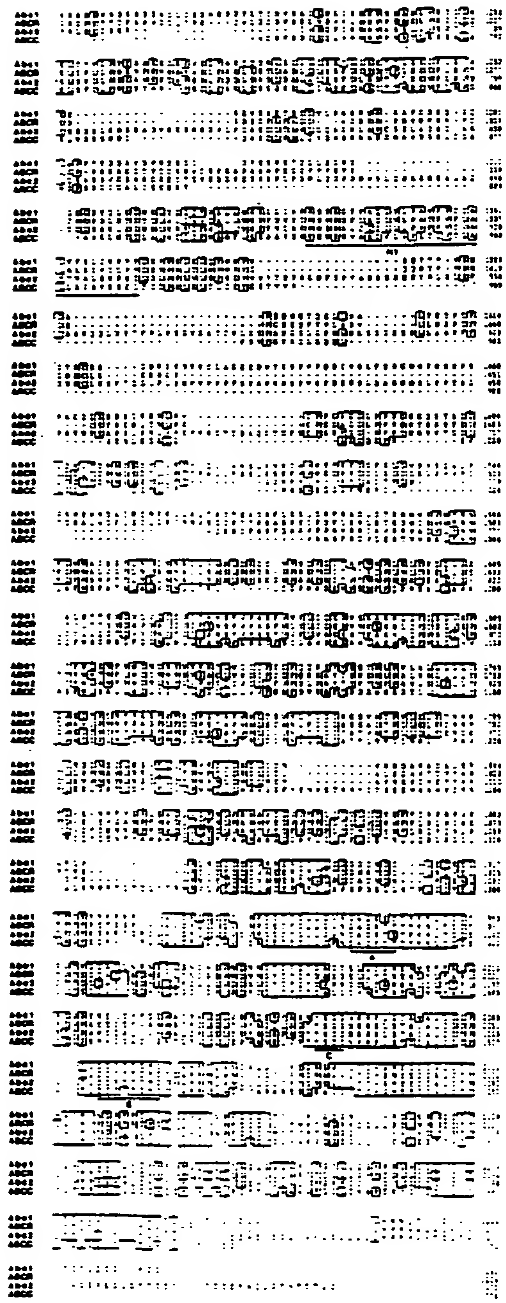


FIGURE 4B

MAPPING THE MOUSE *ABCR* LOCUS

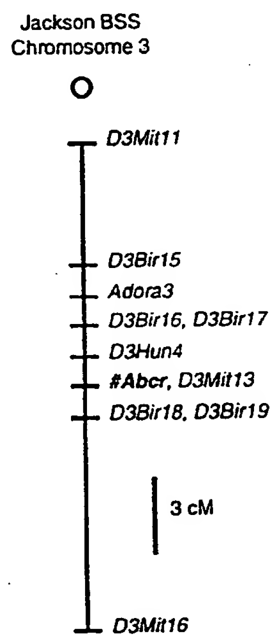
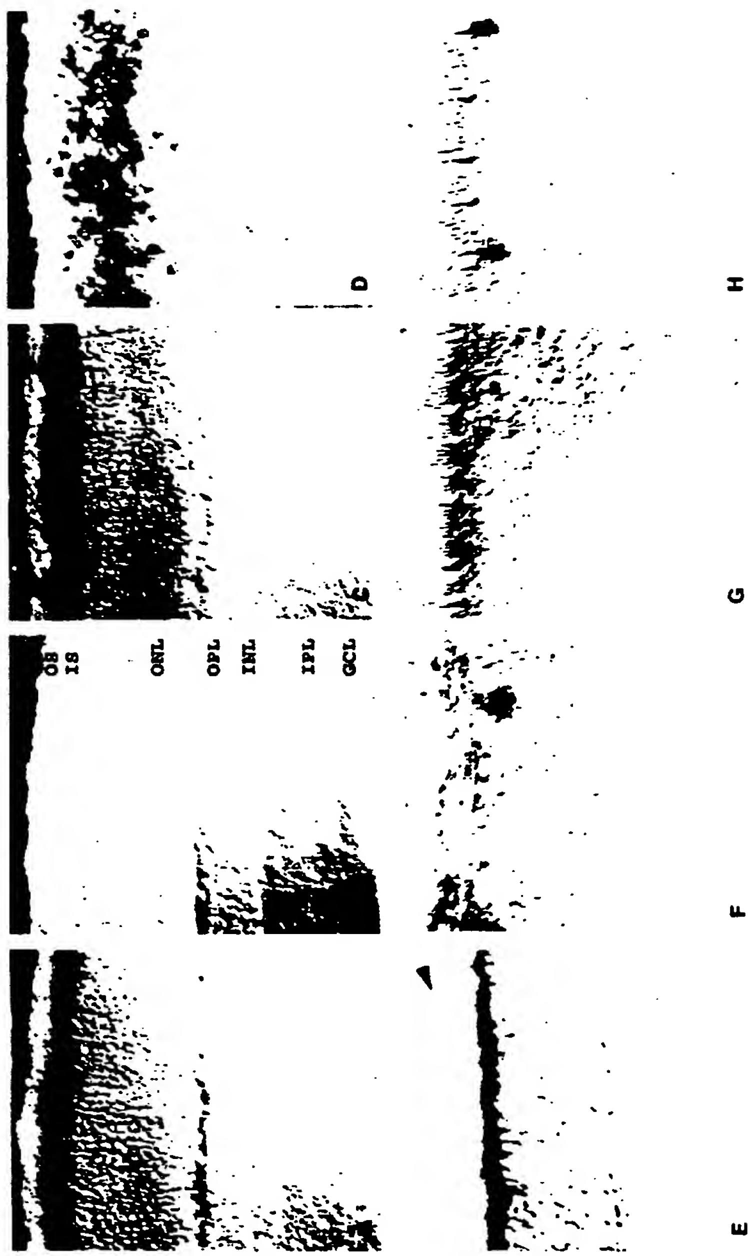


Figure 5

Figure 6



ABCR patent
Figure 7

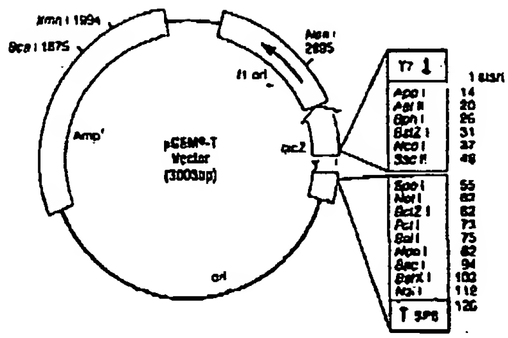

Bca I 1875
Amp
f1 ori
Vector
(3000bp)
ori
T7
1
14
20
26
31
37
46
55
62
62
73
75
82
94
103
112
126
SP6

Figure 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/03895

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :435/172.3, 243, 320.1, 325, 410; 514/44; 536/44
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/172.3, 243, 320.1, 325, 410; 514/44; 536/44

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y,P	GERBER, S. et al. Complete exon-intron structure of the retina-specific ATP binding transporter gene (ABCR) allows identification of novel mutations underlying stargardt disease. Genomics. 1998, Vol. 48, pages 139-142, see entire document.	1-3, 6-21, 27, and 55-58
Y	HOYNG, C.B. et al. Genetic fine mapping of the gene for recessive Stargardt disease. Human Genetics. 1996, Vol. 98, pages 500-504, see entire document.	1-3, 6-21, 27, and 55-58
Y,P	SUN, H. et al. Stargardt's ABCR is localized to the disc membrane of retinal rod outer segments. Nature Genetics. 17 September 1997, Vol. 17, pages 15-16, see entire document.	1-3, 6-21, 27, and 55-58

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 JUNE 1998	16.07.98

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	BRIAN R. STANTON Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/03895

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 2, 3, 7-10, 12-15, 17-21, 27, and 55-58 (each in part)
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

The claims have been searched in part only because no computer readable form of the claimed sequences has been submitted. Therefore, the claims have only been searched on text based criteria.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-3, 6-21, 27, and 55-58

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/03895

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

A01N 43/04; C07H 21/02, 21/04; C12N 5/10, 15/00, 15/09, 15/11, 15/12, 15/63, 15/70, 15/74, 15/79

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

Databases: agricola; aidsline; anabstr; aquasci; biobusiness; biosis; biotechabs; biotechds; caba; cancerlit; caplus; ceaba; cen; cin; cjacs; confsci; cropb; cropu; ddfb; dgene; dissabs; drugb; druglaunch; drugnl; drugu; embal; embase; fsta; genbank; healsafe; ifipat; jicst-eplus; kosmet; lifesci; medline; nioshtic; ntis; ocean; phar; phic; phin; promt; scisearch; toxline; toxlit; uspatfull; wpids; APS

Search Terms: retina?; specific?; atp; adenosin?; bind?; transport?; stargardt?; abcr; anderson?/au; allikmets?/au; dean?/au; leppart?/au; lewis?/au; lewis?/au; li y?/au lupski?/au; nathans?;/au; rattner?/au; shroyer?/au; singh?/au; smallwood?/au; sun h?/au

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-3, 6-21, 27, and 55-58, drawn to nucleic acids encoding retina-specific ATP Binding cassette transporter, and methods of using such a nucleic acid.

Group II, claim(s) 4, 5, 22-26, and 29-34, drawn to retina-specific ATP binding cassette transporter proteins and methods of using such proteins.

Group III, claim(s) 28, drawn to antisense nucleic acids.

Group IV, claims 35, 36, 38, and 39, drawn to transgenic animals comprising sequences that encode retina-specific ATP Binding cassette transporter proteins.

Group V, claim 37, drawn to transgenic animals that lack expression of retina-specific ATP Binding cassette transporter protein.

Group VI, claims 40-54, drawn to diagnostic kits comprising primer pairs.

The inventions listed as Groups I-VI do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The unifying technical feature of the invention of group I is a nucleic acid that encodes retina-specific ATP Binding cassette transporter proteins. Such a nucleic acid may be used directly as a pharmaceutical as evidenced by the invention of claim 27 and therefore is, in and of itself useful in the absence of its encoded protein. Moreover, the isolation of such nucleic acids appears to be suggested by Allikmets et al. (1996, abstract) and thus the isolated nucleic acid does not constitute a special technical feature within the meaning of PCT Rule 13.2. In contrast, the special technical feature of the invention of group II is a retina-specific ATP Binding cassette transporter protein which may be used as a tool for screening for agents that alter the protein activity of such a protein (see e.g. claim 29) and therefore does not have the same special technical feature as the nucleic acids of the invention of group I. The special technical feature of the invention of group III is an antisense nucleic acid which inhibits gene expression and therefore has a separate feature than that of the nucleic acids of group I and the proteins of group II. The special technical feature of the invention of group IV is a multicellular animal that has been altered by gain of function by virtue of having a nucleic acid encoding retina-specific ATP Binding cassette transporter incorporated therein. Therefore, this feature is based on the alteration of a multicellular organism which is distinct from an isolated nucleic acid of any type (e.g. groups I and III) or proteins (group II). The special technical feature of the invention of group V is a loss of function of a gene and/or gene product which is distinct from animals that have gain of function because the feature is based on a lack of functionality within an organism. The special technical feature of the invention of group VI is primer pairs

that are useful for the detection of various disease conditions and therefore relates to elaboration of the presence or absence of particular genes and alleles rather than the use of any particular nucleic acid, protein or animal.

Consequently, given the differences among the special technical features indicates that the several inventions are not so linked by any special technical feature within the meaning of PCT Rule 13.2 such they form a single inventive concept as defined by PCT Rule 13.1.

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y,P	ALLIKMETS, R. et al. A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is mutated in recessive Stargardt macular dystrophy. Nature Genetics. 15 March 1997, Vol. 15, pages 236-245, see entire document.	1-3, 6-21, 27, and 55-58